

Dissertation zur Erlangung des Doktorgrades
der Fakultät Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

**Studien zur Untersuchung der Proteine
Rad14/XPA und RPA in der
Nukleotidexzisionsreparatur**

Nina Simon (geb. Wichlein)

aus

Achim (Deutschland)

2017

Erklärung

Diese Dissertation wurde im Sinne von §7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Thomas Carell betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig ohne unerlaubte Hilfe erarbeitet.

München, den 06.11.2017

Nina Simon

Dissertation eingereicht am 10.10.2017

1. Gutachter: *Prof. Dr. Thomas Carell*
2. Gutachter: *PD. Dr. Dietmar Martin*

Mündliche Prüfung am 03.11.2017

Für meine Mama

„Das Schönste, was wir entdecken können, ist das Geheimnisvolle.“

Albert Einstein

Danksagung

An dieser Stelle möchte ich mich bei zahlreichen Personen für Ihre fachliche und persönliche Unterstützung und für ihren Rückhalt während meiner Doktorarbeit ganz herzlich bedanken. Zuerst einmal möchte ich mich bei meinem Doktorvater *Prof. Dr. Thomas Carell* für das herausfordernde, spannende Thema und die exzellenten Arbeitsbedingungen bedanken, sowie für die wissenschaftlichen Konferenzen.

Auch bedanke ich mich bei Prof. Dr. Dietmar Martin für die Übernahme des Koreferats, sowie allen Mitgliedern der Prüfungskommission.

Weiter danke ich *Dr. Markus Müller, Angie Kirchner, Johanna Bretzler, Franziska Traube* und *Michael Stadlmeier* für das sorgfältige Korrekturlesen. Frau *Slava Gärtner* danke ich für ihre Hilfe in allen organisatorischen Angelegenheiten. *Kristof Hufnagel* und *Kerstin Kurz* danke ich für alle organisatorische Labor-Angelegenheiten, besonders für das Gele gießen. Danke *Kristof!*

Dr. Markus Müller danke ich für die Unterstützung bei allen biochemischen und IT-Fragen. *Charlotte Ebert* und *Johanna Bretzler* danke ich für die tolle Zusammenarbeit.

Meinen Labornachbarn *Sebastian Thallmair* danke ich für sein offenes Ohr und seine Hilfsbereitschaft. An dieser Stelle danke ich auch *Franziska Traube*. Ohne sie wären die frühmorgendlichen Tage und Wochenenden im Labor einsam gewesen. Dem ganzen Arbeitskreis, besonders mein Labor F4.045, danke ich für die tolle und freundschaftliche Atmosphäre. Zudem möchte ich mich bei meinen Kooperationspartnern bedanken. Ein großer Dank gilt *Dr. Sabine Schneider* für das Messen unserer Kristalle und die anschließende Auswertung. Hervorheben möchte ich auch *Prof. Caroline Kisker* und *Dr. Jochen Kuper* für ihre wissenschaftliche Unterstützung in unserem XPA Projekt, insbesondere *Jochen* für die vielen Skype Diskussionen.

Meinen Praktikanten möchte ich für die tolle Unterstützung bedanken. Hervorheben möchte ich *Jana Bogena*. Danke, dass du in dieser Zeit nicht nur zu einer meiner besten Freundinnen geworden bist, sondern auch für deine Unterstützung und das du immer für mich da bist.

Ein großer Dank an meine Kochgruppe *Angie Kirchner, Johanna Bretzler, Franziska Traube, Olga Kameroner, Nada Raddaoui, Christina Schneider* und *Miriam Wallner* sowohl für die schönen Kochstunden und vielen leckeren Gerichten als auch für die schöne Zeit (Musical, Kino, Brunchen,...). Während meiner Dissertation wurde ich von so tollen Menschen begleitet und unterstützt. Danke *Charlotte Sievers, Alexandra Reimann, Jessica Pongers, Svenja Laue* und *Jana Umbach* für eure Freundschaft. Ohne euch wäre die Zeit nicht so toll gewesen!

Zum Schluss gilt mein herzlichster Dank meiner Familie, die mich immer unterstützt und mir Halt gaben. Danke *Mama, Mathias, Sarah* und *Fenja*, dass ihr immer für mich da seid. Zuletzt danke ich

meinen Mann *Martin* vor allem für seine Geduld wenn ich mal wieder am Wochenende oder sehr lange im Labor war, für sein Verständnis und seine Unterstützung.

Publikationsliste

Teile dieser Dissertation wurden bereits publiziert oder auf Konferenzen vorgestellt.

Publikationen:

1. C. Ebert*, N. Simon*, S. Schneider, T. Carell, *ChemBioChem*, **2017**, 18, 1379-1382, „*Structural insights into the recognition of N²-aryl- and C8-aryl DNA lesions by the repair protein XPA/Rad14*”.
2. M. Helm, J. Bretzler, N. Simon, T. Carell, Unendliche Weiten-Kreuz und quer durchs Chemie-Universum, Kapitel 4 (Der erweiterte genetische Code: DNA-und RNA-Basen jenseits von Watson und Crick), 44-49, Wiley-VCH, Weinheim, **2017**.
3. N. Simon*, C. Ebert*, S. Schneider, *Chemistry*, **2016**, 22, 10782–10785, „*Structural Basis for Bulky-Adduct DNA-Lesion Recognition by the Nucleotide Excision Repair Protein Rad14*”
4. S. C. Koch, N. Simon, C. Ebert, T. Carell, *Quarterly Reviews of Biophysics*, **2016**, 49, e5, „*Molecular mechanisms of xeroderma pigmentosum (XP) proteins*”.
5. S. C. Koch*, J. Kuper*, K. L. Gasteiger*, N. Simon, R. Strasser, D. Eisen, S. Geiger, S. Schneider, C. Kisker, T. Carell, *PNAS* **2015**, 112, 8272–8277, „*Structural insights into the recognition of cisplatin and AAF-dG lesion by Rad14 (XPA)*”.

* Beitrag zu gleichen Teilen.

Konferenzbeiträge:

1. 1st PSL Chemical Biology Symposium in Paris, Frankreich (2016) – Posterpräsentation: „*Interaction of XPA and RPA in the Nucleotide Excision Repair*”.
2. SFB 749 Meeting in Kloster Irsee, Deutschland (2016) – Posterpräsentation: „*Investigations towards a specific and reversible protein marker using pyridoxal phosphate derivatives*”.
3. Internes Gruppenseminar von Prof. Dr. Carell in Venedig, Italien (2015) – Vortrag: „*Cryo-Electron Microscopy*”.
4. Gordon Research Konferenz „Nucleosides, Nucleotides & Oligonucleotides” in Newport RI, USA (2015) – Posterpräsentation: „*XPA as a damage sensor in Nucleotide Excision Repair*”.
5. SFB 749 Meeting, München, Deutschland (2014) – Posterpräsentation: „*DNA damage recognition by XPA/Rad14 in Nucleotide Excision Repair*”.

Inhaltsverzeichnis

Zusammenfassung	VI
Summary	X
1. Einleitung	1
1.1. Molekulare Mechanismen der <i>Xeroderma Pigmentosum</i> (XP) Proteine	2
1.2. <i>Bulky</i> DNA-Addukte - Die Substrate der Nukleotidexzisionsreparatur	35
1.3. Replikationsprotein A	39
1.3.1. Struktureller Aufbau von RPA	39
1.3.2. RPA-DNA Bindung	40
1.3.3. Die verschiedenen Funktionen von RPA	41
1.3.4. Regulation von RPA – Allosterie.....	44
1.3.5. Studien zur Strukturaufklärung von RPA	44
1.3.5.1. Elektronenmikroskopie von RPA.....	44
1.3.5.2. NMR Strukturen von RPA.....	45
1.3.5.3. Kristallstrukturen von RPA	47
1.4. XPA und RPA	51
1.4.1. Interaktion von XPA und RPA.....	51
1.4.2. Studien von XPA und RPA	52
2. Aufgabenstellung	55
3. Das Hefeprotein Rad14	57
3.1. Struktureller Einblick in die Erkennung von Cisplatin- und AAF-dG Schäden durch Rad14 (XPA)	57
3.1.1. Einführung.....	57
3.1.2. Autorenbeitrag.....	57
3.2. Strukturelle Basis für <i>bulky</i> DNA-Adduktschadenserkennung durch das Nukleotidexzisionsreparaturprotein Rad14	71
3.2.1. Lizenz.....	71
3.2.2. Einführung.....	71
3.2.3. Autorenbeitrag.....	71
3.3. Strukturelle Einblicke in die Erkennung von N^2 -Aryl- und C8-Aryl DNA-Schäden durch das Reparaturprotein XPA/Rad14	103
3.3.1. Lizenz.....	103
3.3.2. Einführung.....	103

3.3.3. Autorenbeitrag.....	103
3.4. Unveröffentlichte Arbeit: <i>In vivo</i> Studien zur Reparatur von AF-Schäden.....	116
3.4.1. Einleitung	116
3.4.2. Projektbeitrag	116
3.4.3. Material und Methoden.....	116
3.4.3.1. Synthese des 18mer Oligonukleotids mit einem C8-AF-dG Schaden	116
3.4.3.1.1. Synthese von C8-AAF-dG	116
3.4.3.1.2. Oligonukleotidsynthese	116
3.4.3.1.3. DNA-Entschützung	117
3.4.3.1.4. Synthese von C8-AF-dG.....	117
3.4.3.1.7. Bestimmung der DNA-Konzentration	119
3.4.3.2. <i>In vivo</i> Studien.....	119
3.4.4. Ergebnisse und Diskussion	119
3.5. Unveröffentlichte Arbeit: Strukturelle Einblicke in die Erkennung von geschädigter DNA mit einer <i>Bubble</i> durch das Reparaturprotein XPA/Rad14	121
3.5.1. Einführung.....	121
3.5.2. Projektbeitrag	121
3.5.3. Material und Methoden.....	121
3.5.3.1. Synthese der Oligonukleotide.....	121
3.5.3.2. Biochemische Studien	122
3.5.3. Ergebnisse	122
4. Unveröffentlichte Arbeiten – Die humanen Proteine XPA und RPA.....	125
4.1. Einleitung und Projektbeiträge zu diesem Projekt	125
4.2. Material und Methoden	126
4.2.1. Material.....	126
4.2.1.1. Chemikalien und Lösungsmittel.....	126
4.2.1.2. Oligonukleotide.....	126
4.2.1.3. Bakterien	128
4.2.1.4. Enzyme.....	128
4.2.1.5. Vektoren.....	129
4.2.1.6. Medien und Puffer	130
4.2.1.6.1. Medien zur Kultivierung von <i>E.coli</i>	130
4.2.1.6.2. Stocklösungen	130
4.2.1.6.3. Agarosegel.....	130

4.2.1.6.4. Native-PAGE.....	131
4.2.1.6.5. SDS-PAGE	131
4.2.1.6.6. XPA-Proteinreinigung mit einer Ni-NTA-Säule.....	133
4.2.1.6.7. XPA-Proteinreinigung mit einer Heparin-Matrix	133
4.2.1.6.8. XPA-Proteinreinigung mit einer Anionenaustaucher-Matrix (MonoQ).....	133
4.2.1.6.9. XPA-Proteinreinigung mit Größenausschlusschromatographie (SEC).....	134
4.2.1.6.10. RPA-Proteinreinigung mit einer GEBLue-Matrix.....	134
4.2.1.6.11. RPA-Proteinreinigung mit einer HAP-Matrix	135
4.2.1.6.12. RPA-Proteinreinigung mit einer Anionenaustaucher-Matrix (MonoQ).....	135
4.2.1.6.13. Kristallisations-, Gelfiltrationsassay- und Kryo-EM-Puffer.....	135
4.2.1.7. Protein- und DNA-Marker	137
4.2.1.8. Präparationskits	137
4.2.1.9. Geräte	138
4.2.2. Methoden	140
4.2.2.1. Mikrobiologische Methoden.....	140
4.2.2.1.1. Bakterienkulturen	140
4.2.2.1.2. Transformation chemisch kompetenter <i>E. Coli</i> Zellen.....	140
4.2.2.1.3. Heterologe Proteinexpression in <i>E.coli</i>	141
4.2.2.1.4. Zellaufschluss mit Ultraschall.....	141
4.2.2.2. Molekularbiologische Methoden.....	142
4.2.2.2.1. DNA Isolierung - Mini-Plasmidpräparation	142
4.2.2.2.2. Klonierung mittels StarGate®-Technologie.....	142
4.2.2.2.3. Restriktionsverdau	142
4.2.2.2.4. DNA-Agarosegelelektrophorese	143
4.2.2.2.5. DNA-Sequenzierung	144
4.2.2.3. Proteinchemische Methoden	144
4.2.2.3.1. XPA _{fl} Reinigung	144
4.2.2.3.2. RPA Reinigung	146
4.2.2.3.3. Bestimmung der Proteinkonzentration	148
4.2.2.3.4. Konzentrierung der Proteinlösung.....	148
4.2.2.3.5. SDS-PAGE	148
4.2.2.5. Biochemische Methoden	149
4.2.2.5.1. EMSA	149
4.2.2.5.2. Gelfiltrationsassay.....	151

4.2.2.5.3. Thermofluor	152
4.2.2.6. Proteinkristallisation	153
4.2.2.7. Negativfärbung und Kryo-Elektronenmikroskopie	154
4.2.2.8. Synthese von Oligonukleotiden mit einem C8-AAF-dG Schaden.....	154
4.2.2.8.1. Synthese von C8-AAF-dG	154
4.2.2.8.2. Oligonukleotidsynthese	155
4.2.2.8.3. DNA-Entschützung	155
4.2.2.8.4. Aufreinigung der Oligonukleotide durch HPLC	155
4.2.2.8.5. MALDI-TOF-Analyse der Oligonukleotide	156
4.2.2.8.6. Bestimmung der DNA-Konzentration	156
4.2.2.8.7. Hybridisierung von DNA-Proben.....	156
4.3. Ergebnisse und Diskussion.....	157
4.3.1. Synthese des C8-AAF- Phosphoramidits und DNA Festphasensynthese	157
4.3.2. Das humane XPA Protein	160
4.3.2.1. Allgemeines.....	160
4.3.2.2. Klonierung.....	160
4.3.2.3. Heterologe Expression und Aufreinigung von XPA.....	161
4.3.3. Das humane Protein RPA	165
4.3.3.1. Allgemeines.....	165
4.3.3.2. Aufreinigung von RPA	165
4.3.4. Isolierung eines stabilen XPA/RPA/DNA-Komplexes	167
4.3.4.1. Allgemeines.....	167
4.3.4.2. Thermofluor Assay	167
4.3.4.3. Gelfiltrationsassay.....	169
4.3.4.3.1. Gelfiltration eines Standards und der einzelnen Proteine.....	169
4.3.4.3.2. Isolierung des XPA/RPA/DNA-Komplexes durch einen Gelfiltrationsassay	171
4.3.4.3.3. Optimierung des Gelfiltrationsassays	173
4.3.4.4. Studien zum XPA/RPA/DNA-Komplex.....	176
4.3.4.5. Screening von DNA-Sequenzen für die Komplexbildung	179
4.3.4.5.1. Die verschiedenen DNA-Substrate.....	179
4.3.4.5.2. DNA Substrat: 34mer mit einem 20nt 5'-Überhang	181
4.3.4.5.3. DNA Substrat: 25mer mit einem 20nt 5'-Überhang	182
4.3.4.5.4. DNA-Substrat: AAF-geschädigte <i>Bubble</i> -DNA.....	184
4.3.5. <i>Crosslinking</i> Studien	189

4.3.6. Kristallisationsstudien	191
4.3.7. Kryo-Elektronenmikroskopie	193
4.3.8. Zusammenfassung	197
5. Ausblick	198
6. Anhang	200
6.1. Gensequenzen XPA.....	200
6.2. Gensequenzen Rad14.....	201
6.3. Gensequenzen RPA.....	201
6.4. Pufferbedingungen für den Thermofluor Assay	203
6.5. Gelfiltrationsassay	204
6.5.1. Retentionsvolumina der einzelnen DNA-Sequenzen	204
6.5.2. Gelfiltrationsassay mit den einzelnen Proteinen und DNA	205
6.5.3. Gelfiltrationsassay mit 15mer als Substrat	208
6.5.4. Gelfiltrationsassay mit 25mer als Substrat	209
6.5.5. Gelfiltrationsassay mit 34mer als Substrat	213
6.5.6. Gelfiltrationsassay mit <i>Bubble</i> -DNA als Substrat	214
6.6. EMSA-Bindungsstudien	217
6.7. Daten der Kristallstrukturen	222
7. Abkürzungsverzeichnis.....	224
8. Literaturverzeichnis	229

Zusammenfassung

Jeden Tag ist unsere Desoxyribonukleinsäure (DNA) schädlichen Einflüssen aus der Umwelt oder aus endogenen Prozessen ausgesetzt. Um die genomische Integrität zu erhalten, haben sich mehrere Reparatursysteme in der Natur entwickelt. Die Nukleotidexzisionsreparatur (NER) spielt eine wichtige Rolle bei der Entfernung von strukturell unterschiedlichen DNA-Schäden, wie sterisch anspruchsvollen Basenaddukten. Bei dieser Reparatur hat das eukaryontische *Xeroderma pigmentosum* A (XPA) Protein eine entscheidende Bedeutung, denn Defekte im XPA Protein führen zu den stärksten Phänotypen von *Xeroderma pigmentosum*. Allerdings wird die genaue Rolle bzw. Funktion von XPA bis heute noch untersucht.^[1]

Daher sollte im ersten Teil dieser Arbeit mehr Kenntnis über den Erkennungsmechanismus beziehungsweise die Rolle von Rad14, das Hefehomolog des humanen XPA Proteins, in der NER erlangt werden. Die dabei entstandene Kristallstruktur mit der DNA-bindenden Domäne von Rad14 und dG(C8)-Acetylaminofluoren (AAF) geschädigter DNA zeigte erste Einblicke in den Bindungsmechanismus von Rad14 an geschädigte DNA (Abbildung 1).

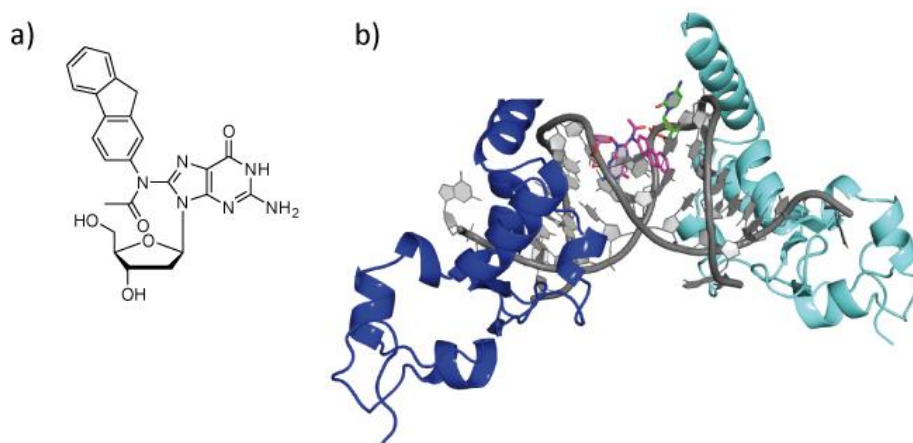


Abbildung 1: a) Darstellung des AAF(C8)-dG Schadens. b) Kristallstruktur von Rad14 mit geschädigter DNA. Die DNA ist in grau, der C8-AAF Schaden in pink und die gegenüber liegende Base dC ist in grün markiert. Die DNA kann von zwei Rad14 Proteinen (in hell- und dunkelblau markiert) gebunden werden. Zudem interkaliert der Schaden in der DNA-Duplex. (PDB: 5A3D).

Diese Kristallstruktur hat eine strukturelle Similarität zur NMR Struktur der DNA-bindenden Domäne von XPA. In dieser nimmt Rad14 eine α/β Architektur, bestehend aus sieben α -Helices und drei β -Faltblättern, ein. Dazu binden zwei Rad14 Proteine oberhalb und unterhalb des Schadens ohne Interaktion mit dem Schaden selbst und bilden dabei ein 13mer Erkennungsmotiv aus. Des Weiteren führt die Bindung des Rad14 Proteins zur Ausbildung eines DNA-Knicks von ungefähr 70° , wobei sich der Schaden in der Knickstelle befindet. Vermutlich kann Rad14 durch einen indirekten Mechanismus Schäden in der DNA erkennen, indem es die Biegsamkeit und Flexibilität der DNA prüft.

Zudem ist der Schaden in dem DNA-Duplex interkaliert und die gegenüberliegende Base wird herausgeklappt, wodurch der Schaden als ein vollständiges Watson-Crick-Basenpaar fungiert. Die dabei möglichen π -stacking Interaktionen könnten für die Stabilisierung des DNA-Knicks und damit einhergehend für die Ausbildung eines stabilen Erkennungskomplexes eine bedeutende Rolle spielen. Um die Rolle der π -stacking Interaktionen näher zu untersuchen, wurden verschiedene Schadensanaloga von AAF synthetisiert, in denen die Größe der aromatischen Fläche verändert wurde. Nach erfolgreicher Synthese von dG(C8)-Acetylphenyl-, dG(C8)-Acetylnaphthyl, dG(C8)-Acetylanthracenyl- und dG(C8)-Acetylpyrenyl-Phosphoramiditen wurden diese durch DNA-Festphasensynthese in die jeweiligen Oligonukleotide eingebaut. Als nächstes wurden die Oligonukleotide mit der DNA-bindenden Domäne von Rad14 kristallisiert. Dabei konnten alle vier Kristallstrukturen mit Rad14 und den unterschiedlichen Schäden erhalten werden, die alle analog zur AAF-Kristallstruktur sind. Auch die unterschiedlichen Basenaddukte sind in die DNA interkaliert und bilden π -stacking Interaktionen aus. Dadurch konnte die Theorie, dass die π -stacking Interaktionen zur Stabilisierung des DNA-Knicks führen, bestätigt werden. Zudem bestätigen die Kristallstrukturen, dass Rad14 die DNA-Schäden in einem *sequence-independent shape-readout* Mechanismus erkennt, indem es die DNA auf Biegsamkeit überprüft und dadurch eine enorm wichtige Rolle in der NER spielt.

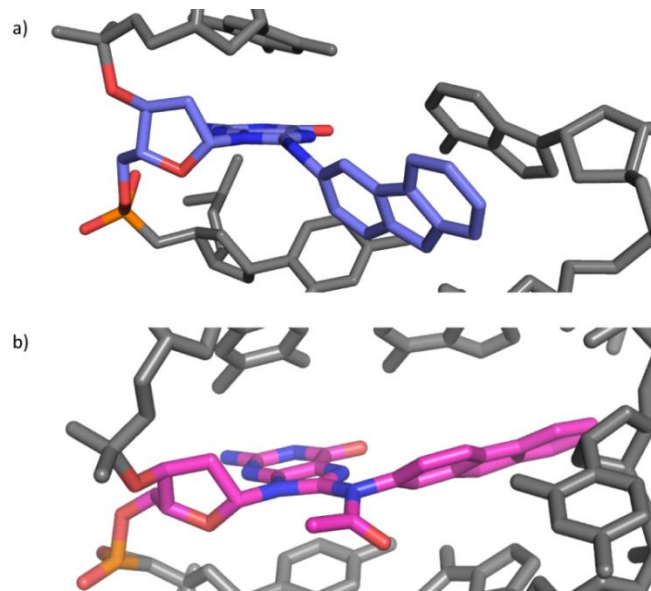


Abbildung 2: Vergleich eines Ausschnitts der Rad14 Kristallstrukturen a) C8-AF-dG (in blau) und b) C8-AAF-dG (in pink). Beide Schäden nehmen eine *syn*-Konformation um die nukleosidische Bindung ein. Im Gegensatz zum C8-AAF Schaden, der mit der dG Base in einer Ebene liegt, stehen die dG Base und der C8-AF Schaden fast senkrecht zueinander. (PDB: 5A3D; 5LCL).

Zusätzlich stellte sich die Frage, wie Rad14 nicht-acetylierte Schäden und vor allem N^2 -geschädigte DNA Addukte erkennt. Dabei konnten die Kristallstrukturen von dG(C8)-AF und dG(N^2)-Acetylaminoanthracen geschädigter DNA mit Rad14 erhalten werden, die auch ähnlich zu den

anderen C8-Schäden sind. Bei genauer Betrachtung der Kristallstrukturen wird erkennbar, dass der AF-Schaden wie AAF eine *syn*-Konformation um die nukleosidische Bindung einnimmt. Der Fluorenring von AF steht fast senkrecht zum Guanosin, wobei AAF in einer Ebene mit Guanosin liegt und AAF somit eine bessere π -Oberfläche ausbildet. Durch die bessere Ausbildung von π -stacking forces von AAF kann der DNA-Knick besser stabilisiert werden und wird daher auch leichter von Rad14 erkannt (Abbildung 2). Im Gegensatz dazu nimmt der N^2 -Schaden eine *anti*-Konformation um die nukleosidische Bindung ein. Auch der Helix-störende Effekt ist geringer als beim AF-Schaden. Dazu liegt das Naphthalin auch in einer Ebene mit Guanosin und bildet eine gute π -Oberfläche aus. Demzufolge ist es für die Schadenserkenkung durch Rad14 wichtig, dass der Schaden den DNA-Knick stabilisieren kann. In *knock-out* Experimenten konnte gezeigt werden, dass C8-AF-geschädigte DNA durch die globale Genomreparatur (GGR) erkannt und entfernt wird. Dennoch werden 60% des Schadens von der RNA-Polymerase überlesen. Dies könnte daran liegen, dass AF eine *syn*- oder *anti*-Konformation einnehmen kann. In der *syn*-Konformation kann es wie AAF erkannt werden und in der *anti*-Konformation kann der Schaden während der Transkription überlesen werden.

Am Ende der Doktorarbeit konnte Rad14 mit einer *Bubble*-DNA, wobei der dG(C8)-AAF Schaden des 15mers in einer 3nt *Bubble* vorliegt, kristallisiert werden. Die Kristallstruktur ist vergleichbar mit den anderen Kristallstrukturen. Allerdings ist die Elektronendichte um den Schaden nicht gut aufgelöst und die Positionen der ungepaarten Basen der Bubble sind auch unzureichend definiert. Aus diesen Ergebnissen lässt sich schließen, dass die π -stacking Interaktionen zwar eine relevante Rolle spielen, aber die Flexibilität der DNA ausschlaggebend für die Erkennung durch Rad14 ist.

Im zweiten Teil sollten weitere Einblicke in den NER-Mechanismus erlangt werden, indem ein stabiler XPA/RPA/DNA-Komplex isoliert und die Struktur mittels Kristallisation oder Kryoelektronenmikroskopie aufgeklärt werden sollte. Das Replikation Protein A (RPA) spielt ebenfalls eine Rolle in der NER und kann sowohl mit XPA interagieren als auch DNA binden.

Dabei konnte ein Gelfiltrationsprotokoll entwickelt werden, welches es ermöglicht den XPA/RPA/DNA-Komplex zu isolieren. Die Protein enthaltenden Fraktionen konnten mittels SDS-PAGE identifiziert und durch EMSA Experimente konnte die Protein-DNA Interaktion bestätigt werden. Außerdem wurden die Pufferbedingungen, Inkubationszeiten, die Protein-DNA Konzentrationsverhältnisse und die DNA Substrate für den Gelfiltrationsassay optimiert. Des Weiteren konnte gezeigt werden, dass der erhaltene Komplex sehr stabil ist. Durch unterschiedliche Zugabezeiten von XPA, RPA oder DNA zum Komplex konnte herausgefunden werden, dass sich die Komplexbildung nicht verändert bei unterschiedlicher Zugabe. Zusätzlich konnten durch den Gelfiltrationsassay verschiedene DNA-Sequenzen für die Komplexbildung überprüft werden, um das am besten geeignete Substrat zu finden. Dabei konnten einige passende DNA-Substrate zur Komplexbildung

gefunden werden. Unter anderem bildeten sowohl die ungeschädigten als auch die AAF-geschädigten DNA-Stränge mit einem Überhang (34mer mit 20nt 5'-Überhang und 25mer mit 20nt 5'-Überhang) mit XPA/RPA einen Komplex aus. Dies lässt vermuten, dass die DNA-Verzweigung, also nicht AAF, schon von XPA/RPA erkannt werden kann. Im Gegensatz zu ungeschädigter *Bubble*-DNA konnte auch mit verschiedenen langen AAF-geschädigten *Bubble*-DNAs gute Komplexbildung erzeugt werden. Somit konnte gezeigt werden, dass die Komplexbildung mit *Bubble*-DNA schadensabhängig ist.

In den Kristallisationsstudien des XPA/RPA/DNA-Komplexes konnten erste Mikrokristalle erhalten werden. Gleichzeitig wurde versucht, die Struktur des Komplexes durch Kryo-Elektronenmikroskopie zu lösen. Dabei konnte durch erste Negativfärbungen eine Partikelgröße von ca. 6-15 nm und durch 2D-Klassifizierung erste strukturelle Hinweise (XPA/RPA mit einem 34mer AAF 20nt 5'-Überhang) erhalten werden. Dennoch konnte bisher keine hochaufgelöste Struktur mittels Kryo-Elektronenmikroskopie gelöst werden. In zukünftigen Untersuchungen müssten daher weitere Strukturstudien durchgeführt werden, um die Struktur des XPA/RPA/DNA-Komplexes zu lösen und somit weitere mechanistische Einblicke in die NER zu bekommen.

Summary

Our deoxyribonucleic acid (DNA) is daily exposed to harmful influences from the environment and endogenous processes. In order to maintain the genomic integrity, several repair systems have been developed in nature. The nucleotide excision repair (NER) plays an important role in the removal of structurally different DNA lesions, such as the bulky base adducts. In this repair, the eukaryotic *Xeroderma pigmentosum* A (XPA) protein is of crucial importance. Defects in the XPA protein can lead to the strongest phenotype of Xeroderma pigmentosum. However, studies to fully understand the exact role or function of XPA are still under active investigation.^[1]

In the first part of this thesis, we investigated the recognition mechanism and the role of Rad14, the yeast homologue of the human XPA protein. The first crystal structure of the DNA-binding domain of Rad14 and dG(C8)-acetylaminofluorene (AAF) damaged DNA provided first insights into the binding mechanism of Rad14 with damaged DNA (Figure 1).

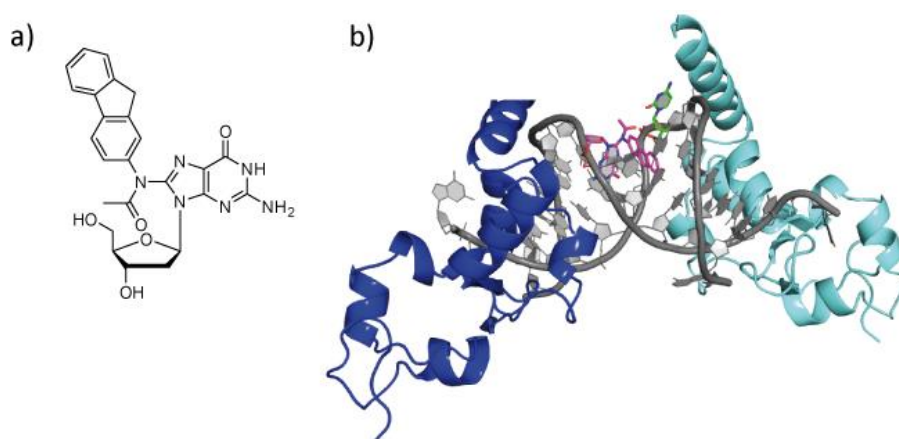


Figure 1: a) Structure of the AAF(C8)-dG lesion. b) Crystal structure of Rad14 with damaged DNA. The DNA is highlighted in grey, the C8-AAF lesion in pink and the opposite base dC in green. Two Rad14 molecules (highlighted in light and dark blue) can bind to the DNA. Additionally, the lesion intercalates into the DNA duplex. (PDB: 5A3D).

This crystal structure has a structural similarity to the NMR structure of the DNA-binding domain of XPA and shows Rad14 in an α/β -architecture consisting of seven α -helices and three β -sheets. Two Rad14 proteins bind above and below the damage to form a 13mer recognition motif. No interaction between the protein and damage was observed. Furthermore, binding of the Rad14 protein leads to the formation of a sharp DNA kink of 70°, with the damage being in the kink. Probably, Rad14 can detect lesions in the DNA through an indirect mechanism by probing the flexibility of the DNA.

Furthermore, the AAF-lesion is intercalated into the DNA duplex and the opposite base is kicked out, causing the lesion to act as a complete Watson-Crick base pair. The possible π -stacking forces could play a significant role in stabilizing the DNA kink. Thus, the investigation of these π -stacking forces is important. Therefore, various lesion analogues of AAF were synthesized in which the size of the

aromatic surface was changed. After successful synthesis of dG(C8)-acetylphenyl-, dG(C8)-acetylnaphthyl-, dG(C8)-acetylanthracenyl- and dG(C8)-acetylpyrenylphosphoramidites, they were incorporated into oligonucleotides by DNA solid-phase synthesis. Next, these oligonucleotides were crystallized with the DNA-binding domain of Rad14. It was possible to obtain all four crystal structures with Rad14 and the lesions which were all similar to the AAF crystal structure. Additionally, we found that all lesions are intercalated into the DNA duplex and form π -stacking forces. These results have strengthened the theory that the π -stacking forces lead to the stabilization of the DNA kink. In addition, the crystal structures confirm that Rad14 recognizes DNA lesions in a sequence-independent shape-readout mechanism by probing the DNA for flexibility and that Rad14 plays an important role in the NER.

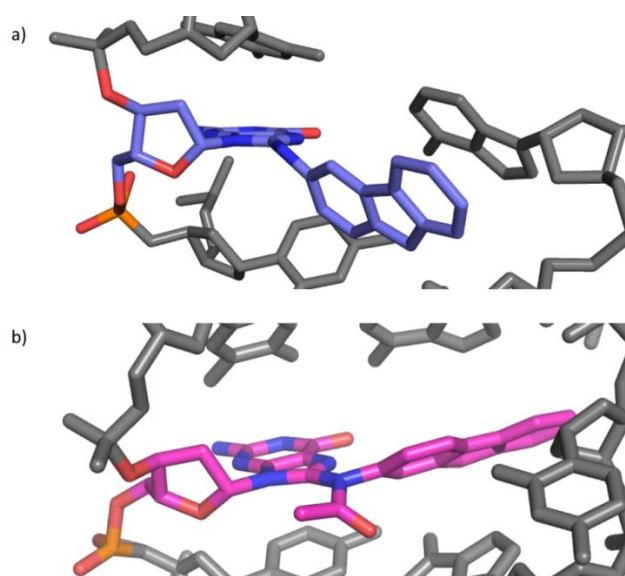


Figure 2: Comparison of a section of the Rad14 crystal structure with a) C8-AF-dG (in blue) and b) C8-AAF-dG (in pink). Both lesions adopt a *syn*-conformation around the nucleosidic bond. The C8-AAF lesion is almost in a plane with the dG base, while the C8-AF lesion and the dG base are perpendicular to each other. (PDB: 5A3D; 5LCL).

We were also wondering, how Rad14 recognizes non-acetylated lesions and N^2 -lesions. The crystal structures of dG(C8)-AF and dG(N^2)-acetylaminonaphthalene-damaged DNA with Rad14 were obtained, which show a similarity to the other C8-lesions. A more detailed analysis of the lesion in the crystal structure revealed that the AF lesion, such as AAF, has a *syn*-conformation around the nucleosidic bond. The fluorene ring of AF is almost perpendicular to the guanosine, while AAF is in plane with guanosine and thus forming a better π surface. Due to the better formation of π -stacking forces of AAF, the DNA kink can be better stabilized and therefore this lesion can be detected more easily by Rad14. In contrast, the N^2 -lesion takes an *anti*-conformation around the nucleosidic bond. The helix disrupting effect is also lower than the AF damage. Further, the naphthalene is in a plane with guanosine and forms a good π surface. Consequently, it is important for damage detection by Rad14 that the damage can stabilize the DNA kink. In knock-out experiments it was shown that the

dG(C8)-AF-lesion is detected and removed by the global genome repair (GGR). Nevertheless, 60% of the lesion is bypassed by the RNA-polymerase. This could be due to the fact that the AF can adopt a *syn*- or *anti*-conformation. In the *syn*-conformation the lesion can be recognized similar to AAF by NER and in the case of the *anti*-conformation the lesion can be bypassed during transcription.

At the end of the thesis, Rad14 was crystallized with a bubble DNA, where the dG(C8)-AAF damage of the 15mer is located in a 3nt bubble. The crystal structure is comparable to the other crystal structures. However, the electron density around the damage is not well resolved and the positions of the mismatch bases in the bubble are inadequately defined. The result indicates that the π -stacking forces play a relevant, but the flexibility of the DNA is crucial for the damage recognition by Rad14.

In the second part we planned to get further insights into the NER mechanism by isolating a stable XPA/RPA/DNA-complex and solving the structure by crystallization or cryo-electron microscopy. Replication protein A (RPA) plays also an important role in the NER and can interact with XPA and binds DNA. A gel filtration protocol was developed, which gives access to the XPA/RPA/DNA-complex. The protein-containing fractions could be identified by SDS-PAGE and the protein-DNA interaction could be confirmed by EMSA experiments. In addition, buffer conditions, incubation times, protein-DNA concentration ratios, and DNA substrates were optimized for the gel filtration assay. Furthermore, it was shown that the resulting complex is very stable. By changing the addition times of XPA, RPA or DNA to the complex it was found that the complex formation doesn't change. Furthermore, different DNA substrates could be screened for the complex formation via gel filtration assay in order to find the most suitable substrate. Some DNA substrates could be identified for complex formation, including the undamaged and AAF damaged DNA strands with overhang (34mer 20nt 5'-overhang and 25mer 20nt 5'-overhang). This suggests that the junction, and not the AAF itself, can already be detected as a lesion for XPA/RPA. In contrast to undamaged bubble DNA, complex formation could also be obtained with different sizes of AAF-damaged bubble DNAs. Thus, it was shown that the complex formation is lesion-dependent. First microcrystals could be obtained in the crystallization studies of the XPA/RPA/DNA-complex. Simultaneously, we tried solving the structure by cryo-electron microscopy. In the first negative stains, a particle size of 6-15 nm could be obtained and by 2D-classification a cherry-similar structure of the complex (XPA/RPA with a 34mer AAF 20nt 5'-overhang) could be obtained. To our dismay, no highly resolved structure has yet been solved by cryo-electron microscopy. In future studies, further structural studies should be carried out in order to solve the structure of the XPA/RPA/DNA-complex and to gain further mechanistic insights into the NER.

1. Einleitung

Die Desoxyribonukleinsäure (DNA) ist der Träger unserer genetischen Erbinformationen. Durch verschiedene exogene und endogene Einflüsse entstehen pro Zelle und pro Tag ca. 10^6 Schäden an der DNA.^[2-3] Diese können zum einen z.B. durch spontane Hydrolyse des Phosphodiesters und der N-glykosidischen Bindung der DNA, durch Replikationsfehler oder durch verschiedene reaktive Metaboliten gebildet werden.^[4-5] Auch können DNA-Schäden durch exogene Einflüsse, wie ultraviolette (UV)-Strahlung oder Umweltmutagene, ausgelöst werden.^[6] Um die Integrität der genetischen Information zu erhalten, ist jeder Organismus mit verschiedenen Reparatursystemen ausgestattet. DNA-Schäden, die nicht repariert werden konnten, können zum einen zur Apoptose oder zur Entstehung von Krankheiten wie Krebs führen.^[7] Neben den Reparatursystemen wie der Nukleotidexzisionsreparatur (NER), Mismatch-Reparatur (MMR), Rekombinationsreparatur, Basenexzisionsreparatur (BER), homologen Rekombination (HR), der nicht-homologen Endverknüpfung (NHE), sowie des Korrekturlesen von DNA-Polymerasen (*proofreading*) gibt es den Schadenstoleranz-Mechanismus (*translesion DNA synthesis*, TLS). Dabei können sogenannte *low-fidelity* Polymerasen (mit einer geringen Genauigkeit) über den Schaden hinweg lesen.^[8-11] Die NER ist ein hocheffizientes Reparatursystem, welches eine große Vielzahl an strukturell unterschiedlicher DNA-Schäden, die die DNA Helix thermodynamisch destabilisieren, erkennt und repariert. Dazu zählen UV-Schäden, oxidative Schäden, sterisch anspruchsvolle (*bulky*) DNA-Addukte, die durch Umweltmutagene verursacht werden, sowie Quervernetzungen der Basen durch Chemotherapeutika, wie Cisplatin.^[10, 12-13] Bei der eukaryotischen NER sind bis zu 30 Proteine beteiligt, wobei die exakte Funktion mancher Proteine noch nicht genau verstanden ist. Eine Dysfunktion der NER führt zu schwerwiegenden Krankheiten wie z.B. *Xeroderma Pigmentosum* (XP). Des Weiteren spielt das *Xeroderma Pigmentosum A* Protein (XPA) eine wichtige Rolle in der NER, denn durch eine Fehlfunktion von XPA können die DNA-Schäden in der NER nicht repariert werden.^[12, 14-15] Zudem wird bis heute noch diskutiert, welche genaue Rolle XPA in der NER spielt.

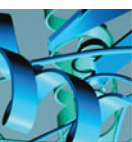
Im folgenden Review wird ausführlich auf die eukaryotische NER sowie auf die einzelnen XP Proteine eingegangen.

1.1. Molekulare Mechanismen der Xeroderma Pigmentosum (XP) Proteine

Sandra C. Koch, Nina Simon, Charlotte Ebert und Thomas Carell, *Quarterly Reviews of Biophysics*, **2016**, 49, e5. „Molecular mechanisms of xeroderma pigmentosum (XP) proteins“. (reproduced with permission).

Lizenz: Kopie der Publikation mit Erlaubnis des Verlags; Copyright 2016 Cambridge University Press.

Die Nukleotidexzisionsreparatur (NER) ist ein vielseitiges und effizientes DNA-Reparatursystem, das für die Entfernung einer Vielzahl von strukturell unterschiedlicher DNA-Schäden verantwortlich ist. Die breite Substratspezifität der NER variiert von DNA-Schäden, die durch Exposition von ultravioletter Strahlung gebildet werden, bis hin zu zahlreichen sperrigen DNA-Addukten, die durch mutagene Umweltchemikalien und zytotoxische Medikamente, die in der Chemotherapie verwendet werden, gebildet werden. Ein Defekt in der NER führt zu schweren Krankheiten wie Xeroderma pigmentosum (XP). Es sind acht XP-Komplementierungsgruppen bekannt, von denen sieben (XPA-XPG) durch Mutationen in Genen verursacht werden, die an der NER beteiligt sind. Das achte Gen, XPV, kodiert für die DNA-Polymerase η , die über DNA-Schäden repliziert, auch translesion synthesis (TLS) genannt. In den vergangenen zehn Jahren haben sich detaillierte, strukturelle Informationen dieser DNA-Reparaturproteine, die in der eukaryotischen NER und TLS involviert sind, ergeben. Diese Strukturen erlauben uns nun, den molekularen Mechanismus der NER und TLS genauer, insbesondere die breite Substratspezifität der NER, zu verstehen. In diesem Review wollen wir die jüngsten Fortschritte in der Schadenserkennung und -reparatur, sowie in der Schadenstoleranz durch die XP-Proteine hervorheben.



REVIEW

Molecular mechanisms of xeroderma pigmentosum (XP) proteins

Sandra C. Koch, Nina Simon, Charlotte Ebert and Thomas Carell*

Center for Integrated Protein Science at the Department of Chemistry, Ludwig-Maximilians Universität München, Butenandstr. 5-13, 81377 Munich, Germany

Quarterly Reviews of Biophysics (2016), 49, e5, pages 1 of 32 doi:10.1017/S0033583515000268

Abstract. Nucleotide excision repair (NER) is a highly versatile and efficient DNA repair process, which is responsible for the removal of a large number of structurally diverse DNA lesions. Its extreme broad substrate specificity ranges from DNA damages formed upon exposure to ultraviolet radiation to numerous bulky DNA adducts induced by mutagenic environmental chemicals and cytotoxic drugs used in chemotherapy. Defective NER leads to serious diseases, such as xeroderma pigmentosum (XP). Eight XP complementation groups are known of which seven (XPA–XPG) are caused by mutations in genes involved in the NER process. The eighth gene, XPV, codes for the DNA polymerase η , which replicates through DNA lesions in a process called translesion synthesis (TLS). Over the past decade, detailed structural information of these DNA repair proteins involved in eukaryotic NER and TLS have emerged. These structures allow us now to understand the molecular mechanism of the NER and TLS processes in quite some detail and we have begun to understand the broad substrate specificity of NER. In this review, we aim to highlight recent advances in the process of damage recognition and repair as well as damage tolerance by the XP proteins.

Key words: NER, xeroderma pigmentosum, crystal structures, DNA repair.

1. Introduction 2

- 1.1. DNA damages 2
- 1.2. NER and NER-related diseases 2
- 1.3. NER mechanism 3
- 1.4. DNA-structure/repair relationships 4
- 1.5. Damage recognition in NER 4
- 1.6. Overview of each XP protein and its role in GG-NER 5

2. XPE (UV-DDB2) 6

- 2.1. Crystal structure of human XPE with 6–4PP damaged dsDNA 7
- 2.2. Crystal structure of human XPE with CPD damaged dsDNA 7
- 2.3. Crystal structure of human XPE with an abasic lesion site in dsDNA 7

3. XPC 7

- 3.1. Crystal structure of yeast XPC (Rad4) in complex with a CPD lesion 9

4. XPB 11

- 4.1. Crystal structure of *Archaeoglobus fulgidus* XPB (AfXPB) 11
- 4.2. Crystal structure of human XPB 12

5. XPD 12

- 5.1. Crystal structure of *Sulfolobus acidocaldarius* XPD (SaXPD) 13

* Authors for correspondence: Thomas Carell, Center for Integrated Protein Science at the Department of Chemistry, Ludwig-Maximilians Universität München, Butenandstr. 5-13, 81377 Munich, Germany. Tel.: +49 (0)89 2180 77755 Email: thomas.carell@cup.uni-muenchen.de



5.2. Crystal structure of <i>T. acidophilum</i> (TaXPD) in complex with ssDNA	13
5.3. Proposed model of damage verification	14
6. XPA	15
6.1. Nuclear magnetic resonance (NMR) structure of the DBD of human XPA	15
6.2. Crystal structures of Rad14 with AAF-dG and cisplatin containing dsDNA	17
7. XPF-ERCC1	17
7.1. Crystal structure of human XPF-ERCC1	18
7.2. NMR structure of C-terminal human XPF in complex with ssDNA	18
8. XPG	19
8.1. Crystal structure of the yeast XPG (Rad2) in complex with dsDNA	19
9. XPV	21
9.1. Crystal structure of the yeast XPV (Rad30) protein	21
9.2. Crystal structure of the yeast XPV (Rad30) in complex with 1,2-GG cisplatin	21
9.3. Crystal structure of the yeast XPV (Rad30) in complex with the CPD lesion	21
9.4. Crystal structure of the human XPV in complex with the CPD lesion	22
10. Conclusion	23
Acknowledgements	24
References	24

1. Introduction

1.1 DNA damages

The DNA is attacked daily by ultraviolet (UV) radiation, chemical carcinogens, cellular metabolites as well as anti-cancer drugs such as cisplatin (Hoeijmakers, 2001) leading to more than 50 000 damages per cell per day (Rajski *et al.* 2000). To preserve the integrity of the genetic information, all organisms are equipped with a number of DNA repair or tolerance mechanisms such as DNA recombination, base excision repair (BER), mismatch repair (MMR), nucleotide excision repair (NER) and translesion synthesis (TLS) that allow cells to cope with the different lesions (Friedberg, 2005; Ohmori *et al.* 2001; Rupp & Howard-Flanders, 1968). In this review, we will focus on eukaryotic NER and TLS.

1.2 NER and NER-related diseases

One of the most prominent processes among the DNA repair pathways is the NER. NER is a central pathway safeguarding the genome and cells against induced carcinogenesis, because of its capacity to eliminate a broad range of structurally very different DNA lesions (de Laat *et al.* 1999; Geacintov *et al.* 2002; Gillet & Scharer, 2006; Gunz *et al.* 1996; Huang *et al.* 1994) including UV lesions such as cyclobutane-pyrimidine dimers (CPDs) and 6-4 pyrimidine-pyrimidone photoproducts (6-4PPs), numerous bulky chemical adducts, intrastrand crosslinks caused by chemotherapeutics such as cisplatin and oxidative damages (Fig. 1) (Brooks *et al.* 2000; Demple & Harrison, 1994). All these lesions do not share the same chemical structures, but they are bulky and they thermodynamically destabilize the DNA helix, except for CPDs.

In humans, mutations of NER components lead to three severe diseases: xeroderma pigmentosum (XP), Cockayne syndrome (CS) and trichothiodystrophy (TTD) (Cleaver, 2000; Lehmann, 2001). The NER pathway is divided in two subpathways: the global genome NER (GG-NER) and the transcription-coupled NER (TC-NER) (Gillet & Scharer, 2006; Hanawalt & Spivak, 2008) differing only in their damage recognition step. But NER in chromatin is a complex process and probably uses the proposed 'access-repair-restore' mechanism, which comprises many components (Scharer, 2013). A defect in TC-NER along with a mild transcription effect can cause CS and TTD (Cleaver *et al.* 2009).

The CS is characterized by developmental impairment, premature aging and sunlight sensitivity. TTD leads to developmental impairment and mental retardation without showing skin cancer predisposition (Cleaver *et al.* 2009; Lehmann, 2003). Mutations inactivating the GG-NER cause the genetic disorder XP (Cleaver *et al.* 2009). Persons affected by XP ('children of the moon') need to avoid exposure to sunlight, because they show a high predisposition to UV-induced skin cancer (Berneburg & Lehmann, 2001; Hebra & Kaposi, 1874; Lehmann, 2003). Seven complementation groups (XPA-XPG)

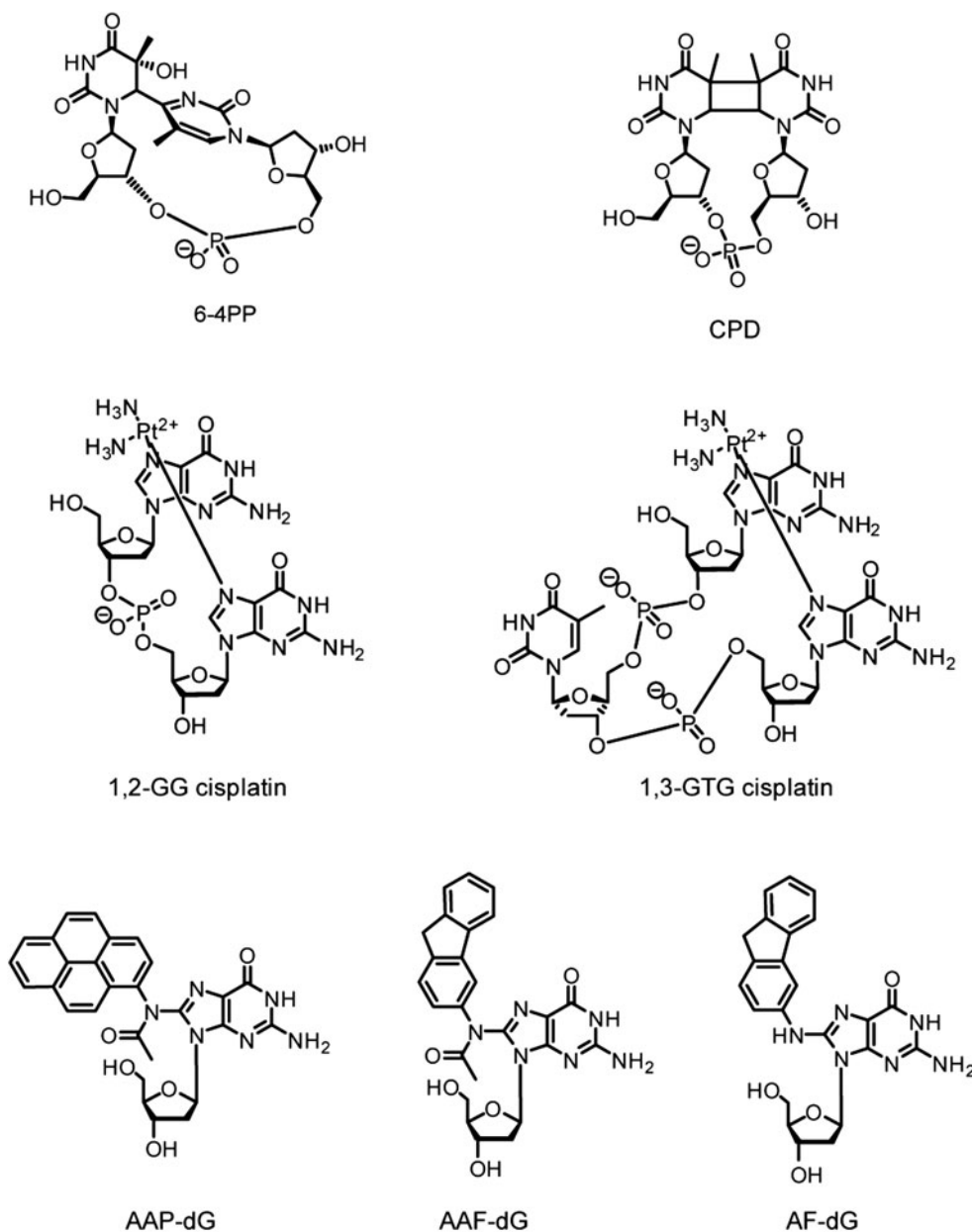


Fig. 1. Examples of DNA adducts repaired by NER. AAF-dG: *N*-(2'-desoxyguanosin-8-yl)-2-acetylaminofluorene, AF-dG: *N*-(2'-desoxyguanosin-8-yl)-2-aminofluorene AAP-dG: *N*-(2'-desoxyguanosin-8-yl)-2-acetylaminopyrene 6-4PP: 6-4PP, CPD: cyclobutane-pyrimidine dimer, 1,2-GG and 1,3-GTG: major cisplatin-induced intrastrand crosslinks lesions.

associated with different NER gene defects were discovered by cell fusion techniques (De Weerd-Kastelein *et al.* 1972). The discovery of NER and its association with genetic disorders together with the mechanism of action have already been reviewed (Friedberg, 1995). Mutations associated with XP and posttranslational modifications of the XP proteins have also been recently reported (Feltes & Bonatto, 2015). The variant form of the human syndrome xeroderma pigmentosum (XPV) (Masutani *et al.* 1999b) is caused by a defect in DNA polymerase η . In this paper, we review the eukaryotic structures of XP proteins (XPA–XPG), including the XP variant DNA polymerase η (called XPV) in complex with damaged or undamaged DNA. This review will provide insights at the atomic level of how lesions in DNA are processed by these proteins.

1.3 NER mechanism

With the exception of chromatin, the NER mechanism is well understood in *Escherichia coli* as well as in eukaryotes (Naegeli & Sugawara, 2011; Sancar & Tang, 1993; Scharer, 2013). While the repair process in eukaryotes is accomplished by up to 30



proteins, the precise role of each protein has not been fully elucidated so far. Despite this, a general core repair mechanism has emerged (Fig. 2). In eukaryotes, the NER is divided in two distinct subpathways differing only in their damage recognition step (de Laat *et al.* 1999; Hanawalt & Spivak, 2008; Mellon *et al.* 1987). TC-NER ensures the repair of lesions which are located on the transcribed strand of active genes (Hanawalt & Spivak, 2008; Hoeijmakers, 2001; Mellon *et al.* 1987; Sweder & Hanawalt, 1992). Here the RNA-polymerase II seems to act as the initial damage sensor. During transcription the polymerase gets stalled at the lesion site (Donahue *et al.* 1994; Tornaletti & Hanawalt, 1999) and by addition of the protein Cockayne syndrome group B (CSB) a stable CSB–RNAP–DNA complex can form by ATP-hydrolysis (Citterio *et al.* 2000; Tantin *et al.* 1997). In case the lesion blocks any read through, downstream factors such as Cockayne syndrome group A protein (CSA) (Kamiuchi *et al.* 2002), TFIIH (Tantin, 1998) and XPG (Iyer *et al.* 1996) are recruited to the damage site and repair is initiated (Hanawalt & Spivak, 2008). The GG-NER in contrast detects and removes helix distorting lesions throughout the whole genome. The exact mechanism of the damage recognition steps is still under debate. Aboussekhra *et al.* and Mu *et al.* were able to reconstitute the repair mechanism *in vitro* allowing deciphering the core excision reaction of NER (Aboussekhra *et al.* 1995; Mu *et al.* 1995). The NER assembly seems to be a sequential and coordinated process in which damage excision and repair is mostly error-free. This repair process is characterized by five main steps (steps I–V): after damage recognition (I) a multi-protein complex ('preincision complex' (PIC)) binds to the damage site (II) leading to a double asymmetric cut of the damage-containing oligonucleotide strand. The damage-containing single-stranded DNA (ssDNA) fragment is eliminated (III). Afterwards the double strand is reconstituted by filling the gap with the DNA polymerase (IV) and ligation of the new synthesized strand takes place (V) (Shivji *et al.* 1995). In this work, we will only consider the GG-NER pathway.

1.4 DNA-structure/repair relationships

Because of its broad substrate specificity, it was speculated early on that the NER system does not recognize the lesion itself but the distortion of the DNA structure. In 1997 Naegeli and co-workers discovered that for damage recognition a chemical modification of the nucleotide in addition to distortion in the DNA helix had to be present (Hess *et al.* 1996b). They could show that neither a modified ribose moiety with intact Watson–Crick base pairing nor a DNA bubble of three base-pairs was recognized and repaired by the NER system. However, placing the modified ribose moiety into a DNA bubble, created a well-recognized NER substrate. These observations helped to formulate the 'bipartite substrate discrimination'-model (Hess *et al.* 1996a, 1997a). This model shows that only DNA damages that cause disrupted base pairing can be recognized as NER substrates. Furthermore, a DNA helix distortion might be additionally needed for the correct recruitment of the NER machinery (Hess *et al.* 1996b). This hypothesis explains why helix-distorting damages are more efficiently processed by NER (Gunz *et al.* 1996; Sugawara *et al.* 2001). These ideas can also explain why 6–4PPs are better substrates than CPD lesions. Not all adducts, induced by polyaromatic hydrocarbons or aromatic amines, are good substrates for NER, because some adducts do not change the duplex structure, e.g. *N*-(deoxyguanosin-8-yl)-2-aminofluorene. The aromatic ring of the lesion is easily accommodated in the minor or major grooves of the DNA duplex. Lesions causing helix disruption upon intercalation of their aromatic ring inside the DNA helix are typically good substrates for NER, e.g. *N*-(deoxyguanosin-8-yl)-2-acetylaminofluorene (dG-AAF), (Gillet *et al.* 2005; Mu *et al.* 1994) or benzo[*a*]pyrene (Buterin *et al.* 2000; Geacintov *et al.* 2002; Hess *et al.* 1997a; Yan *et al.* 2003). Nowadays, it is assumed that all parameters able to reduce the thermodynamic stability of the DNA play a role for the NER damage recognition step. The repair rate is apparently proportional to the decrease of thermodynamic stability of the duplex induced by the discrete DNA lesions (Aboussekhra *et al.* 1995; Geacintov *et al.* 2002; Gillet & Scharer, 2006; Gunz *et al.* 1996).

1.5 Damage recognition in NER

How does discrimination between undamaged and damaged DNA take place in the context of dominating undamaged DNA present in the nucleus? The exact damage recognition process in NER is still unclear and as such it is the most challenging current question associated with the NER process.

Several experiments with fluorescein-labeled proteins, such as ERCC1 (Houtsmuller *et al.* 1999), XPB (Hoogstraten *et al.* 2002) and XPA (Rademakers *et al.* 2003), in combination with photobleaching techniques showed that the repair proteins move freely in the nucleus and accumulate successively at the damaged site. These observations argue that recognition factors must initially bind to the damage site and subsequently recruit downstream factors. At the beginning of NER research XPA was thought to be the initial damage recognizer in NER (Asahina *et al.* 1994; Jones & Wood, 1993). However, in 1998 Sugawara *et al.* (Sugawara *et al.* 1998) demonstrated in an *in vitro* experiment, that the complex XPC–RAD23B was the first damage binder followed by participation of XPA. While for XPC–RAD23B the involvement in the GG-NER pathway has been clearly shown, the complex is not necessary for TC-NER. XPA in contrast, is indispensable for both GG-NER and TC-NER, which perhaps supports its central function in a later step of NER (Kobayashi *et al.* 1998; Sugawara *et al.* 1998). The role of XPA requires further investigations.

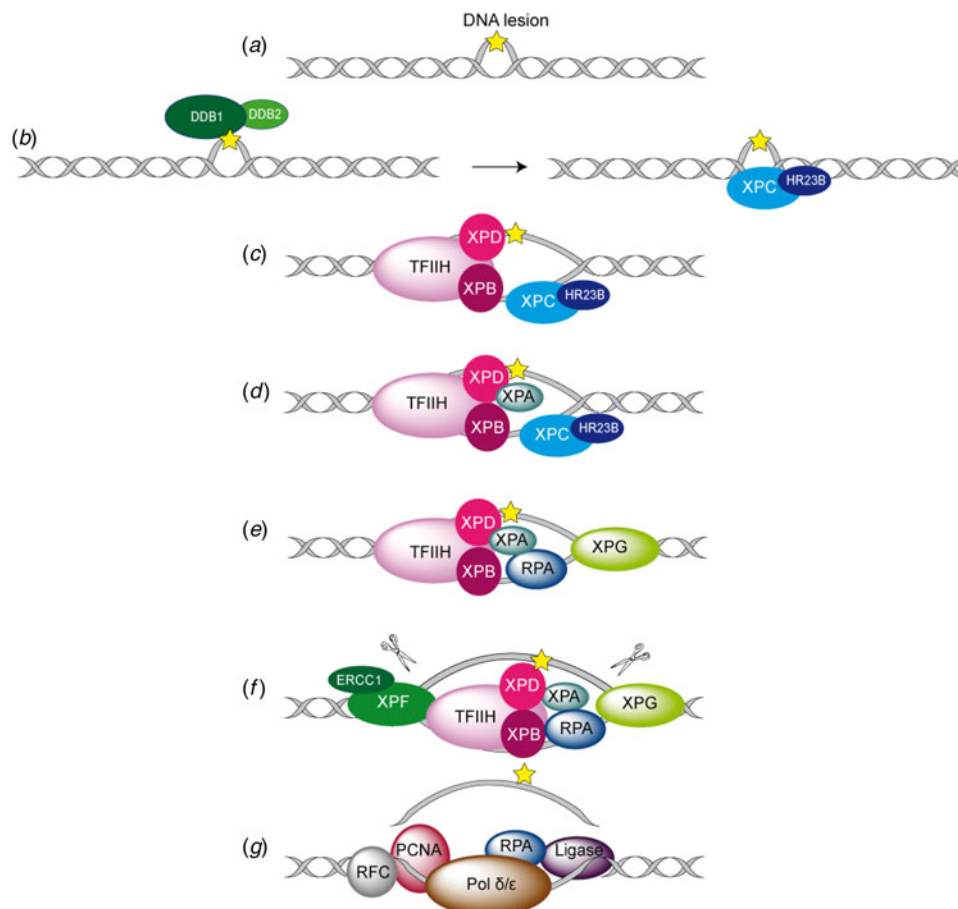


Fig. 2. Overview of the global genome-NER mechanism (Friedberg, 2001; Li *et al.* 2015).

Biochemical experiments demonstrated that the UV-damaged DNA-binding protein (UV-DDB) (Fujiwara *et al.* 1999; Keeney *et al.* 1993; Reardon *et al.* 1993), XPC–RAD23B (Batty *et al.* 2000a; Raoul *et al.* 1995; Sugawara *et al.* 1998, 2001), XPA (Buschta-Hedayat *et al.* 1999; Jones & Wood, 1993; Li *et al.* 1995a) and replication protein A (RPA) (Burns *et al.* 1996; Mellon *et al.* 1987; Reardon & Sancar, 2002; Schweizer *et al.* 1999) show all affinity to damage-containing DNA. XPC–RAD23B binds specifically to DNA lesions such as 6–4PP (Batty *et al.* 2000a; Kusumoto *et al.* 2001; Raoul *et al.* 1995), cisplatin (Evans *et al.* 1997b; Hey *et al.* 2002; Moggs *et al.* 1996, 1997; Sugawara *et al.* 1998), dG-AAF (Sugawara *et al.* 1998) and cholesterol adducts (Kusumoto *et al.* 2001; Sugawara *et al.* 2001). It recognizes CPD damages (Kusumoto *et al.* 2001) poorly although their repair is strongly dependent on the presence of XPC (Venema *et al.* 1991). The affinity of XPC–RAD23B to CPDs can be enhanced by mismatches surrounding the lesion (Sugawara *et al.* 2001). XPA alone or in combination with RPA binds to DNA containing bulky adducts such as dG-AAF (Liu *et al.* 2005; Yang *et al.* 2002). By far the strongest DNA binding for XPA are DNA junctions which suggest that XPA interacts with an intermediate structure rather than directly with a lesion (Missura *et al.* 2001). Photoaffinity-crosslinking experiments demonstrate that XPC–RAD23B (Maltseva *et al.* 2007), as well as XPA (Wakasugi & Sancar, 1999) show a high affinity to damaged DNA without binding directly to the lesion (Reardon & Sancar, 2002). Reconstitution experiments with NER-deficient cells with all the different NER proteins in different order delivered contradictory results. Sugawara *et al.* showed a fast repair rate when the cells were preincubated with XPC–RAD23B (Sugawara *et al.* 1998). Preincubation of the damaged plasmids with XPA–RPA slowed down the repair velocity. Experiments from Wakasugi *et al.* however showed a more effective repair when damaged DNA was preincubated with XPA/RPA (Wakasugi & Sancar, 1999).

1.6 Overview of each XP protein and its role in GG-NER

In Fig. 2, the different steps of the GG-NER mechanism are depicted. It is believed that XPC–RAD23B initiates the GR-NER by recognition of the damage (Fig. 2a, b) (Reardon & Sancar, 2003; Sugawara *et al.* 1998; Wakasugi & Sancar, 1998). XPC ignores the lesion but binds to the ssDNA opposite to the lesion (Min & Pavletich, 2007). In the case of UV lesions, the

Table 1. Overview of the XP proteins involved in NER and TLS. ERCCX: Excision repair cross-complementing rodent repair deficiency complementation group X. Af: *Archaeoglobus fulgidus*, Dr: *Danio rerio*, Sa: *Sulfolobus acidocaldarius*, Ape: *Aeropyrum pernix*, Ta: *Thermoplasma acidophilum*, AAF-dG: N-(2'-deoxyguanosin-8-yl)-2-acetylaminofluorene, 6-4PP: 6-4PP, CPD: cyclobutane-pyrimidine dimer. PDB codes depicted in bold refer to NMR structures.

Protein (gene)	Function	Protein size (aa)	Interaction partner	PDB code	DNA substrate
XPA (XPA)	Damage verification	31 kDa (273)	RPA, TFIIH, XPF-ERCC1, DDB2, PCNA	1XPA 5A39 (Rad14) 5A3D (Rad14)	– cisplatin AAF-dG
XPB (ERCC3)	3'–5' DNA helicase	89 kDa (782)	XPC, XPA, XPG, XPF	2FWR (Af) 4ERN	– –
XPC–RAD23B (XPC)	Damage recognition	XPC: 106 kDa (940) RAD23B: 43 kDa (409)	XPA, TFIIH, RAD23B, UV-DDB2,	2QSG (Rad4)	CPD
XPD (ERCC2)	5'–3' DNA helicase	87 kDa (760)	XPC, XPA, XPG, XPF	3CRV (Sa) 4A15 (Ta)	– ssDNA
XPE (UV-DDB1/2)	Damage recognition of UV lesions	DDB1: 127 kDa (1149) DDB2: 48 kDa (428)	XPC, RPA, XPA	3EI1 (Hs-Dr) 4A08 4E5Z	6-4PP CPD abasic lesion
XPF–ERCC1 (ERCC4)	5'-endonuclease	104 kDa (916) 33 kDa (297)	TFIIH, XPA, ERCC1	2A1J 2KN7	– ssDNA
XPG (ERCC5)	3'-endonuclease	133 kDa (1186)	XPA, TFIIH, PCNA, XPD	4Q0W (Rad2)	dsDNA with ss overhang
XPV (POLH)	Damage bypass	78 kDa (713)	PCNA	1JIH (Rad30) 2R8K (Rad30) 3MFI (Rad30) 3MR3	– cisplatin CPD CPD

recognition process by XPC–RAD23B is sustained by the UV-DNA damage-binding protein 2 (UV-DDB2) (Luijsterburg *et al.* 2010; Min & Pavletich, 2007; Riedl *et al.* 2003; Volker *et al.* 2001). After this step, the general transcription factor TFIIH is recruited to the lesion site using the energy provided by the ATPase activity of XPB (Fig. 2c). Normally, TFIIH is involved in transcription by RNA polymerases I and II, but in the presence of DNA damage it is clearly recruited 5' to the damage site (Hoogstraten *et al.* 2002). Two subunits of TFIIH, the ATP-dependent DNA helicases XPB and XPD, are needed for 'hooking up' TFIIH at the lesion site performing further unwinding and damage verification. It is believed that XPB opens the DNA and facilitates the binding of XPD to the DNA which verifies the chemical modification of the lesion (Scharer, 2013). In addition, XPD scans the strand with the lesion and both helicases are stalled when XPD encounters a bulky lesion (Li *et al.* 2015; Mathieu *et al.* 2013). In the next step, XPA and RPA (Fig. 2d, e) are thought to be recruited to the damaged site to form the PIC which is stabilized by the structure specific endonucleases XPG (Fig. 2e). Upon arrival of XPG to the PIC, XPC–RAD23B leaves the complex (Riedl *et al.* 2003; Wakasugi & Sancar, 1998). Besides XPD, XPA might be involved in damage verification, however, the details of this event remain elusive (Naegeli & Sugawara, 2011; Tapias *et al.* 2004). The ssDNA-binding protein RPA now binds the undamaged strand, while TFIIH opens the duplex. XPA is then thought to recruit the structure specific endonuclease XPF–ERCC1 heterodimer (Li *et al.* 1995a; Tsodikov *et al.* 2007) and incision on both sites of the lesion takes place. The first incision 5' of the lesion is made by XPF–ERCC1 (Staresincic *et al.* 2009). Then, XPG cuts 3' to the lesion (Fagbemi *et al.* 2011) triggering the elimination of a 24–32 oligonucleotide containing the lesion (Fig. 2f) (Edenberg & Hanawalt, 1972; Evans *et al.* 1997b; Ikegami *et al.* 1998). After that, DNA replication by the polymerases $\kappa/\delta/\epsilon$ takes place sustained by other factors (PCNA, RFC and RPA) (Ogi *et al.* 2010). Finally, the new synthesized DNA strand is mostly sealed by DNA ligase III/XRCC1. DNA ligase I is only recruited to sites of UV damage in proliferating cells (Fig. 2g) (Moser *et al.* 2001, 2007; Reardon & Sancar, 2005; Wood, 1997). Table 1 gives an overview of the XP proteins, their function and size, their interaction partners as well as the PDB codes of solved structures associated to the DNA substrates.

2. XPE (UV-DDB2)

For the recognition of UV lesions such as 6-4PPs and CPD, the UV-DDB protein seems to promote binding of XPC–RAD23B to the damaged site (Scrima *et al.* 2008). The UV-DDB protein, consists of the two subunits DDB1 (also called p127 with 127 kDa, 1149 amino acids) and DDB2 (called p48 with 48 kDa, 428 amino acids) (Fig. 3a). It shows a 100- to 1000-fold higher affinity to photolesions than XPC–RAD23B (Keeney *et al.* 1993; Takao *et al.* 1993). Both subunits are required for the activity of the complex (Hwang *et al.* 1998a; Keeney *et al.* 1993; Nichols *et al.* 1996). UV-DDB is known to be involved in the DNA damage recognition process of GG-NER (Wakasugi *et al.* 2002). The DDB2 subunit binds with



preference to 6–4PPs but it also recognizes CPD lesions and it also binds abasic sites as well as DNA bubbles (Fujiwara *et al.* 1999). DDB1 is not involved in DNA binding. Mutation of the DDB2 subunit leads to a loss of function of the UV-DDB protein found in XP complementation group E (XPE) in patients (Chu & Chang, 1988). XPE-deficient cells are shown to be defective in GG-NER (Hwang *et al.* 1999). Wakasugi *et al.* showed that UV-DDB rapidly accumulates *in vivo* independently of XPA and XPC at DNA damage sites after UV-irradiation (Wakasugi *et al.* 2002). Furthermore, they showed that UV lesions are mostly recognized by UV-DDB which then recruits XPA and XPC.

2.1 Crystal structure of human XPE with 6–4PP damaged dsDNA

Scrima *et al.* (2008) solved the crystal structure of the UV-DDB (human DDB1, HsDDB1 and the zebrafish DDB2 ortholog, DrDDB2) protein in complex with double-stranded DNA (dsDNA) containing the 6–4PP photolesion (Fig. 3b). In order to achieve damage recognition the UV-DDB2 protein inserts a β -hairpin into the minor groove of the DNA. In addition, the damaged dinucleotide is flipped out of the double helix and inserted into a binding pocket of the protein leading to a DNA kink of about 40° at the lesion site. This destabilized conformation of the DNA might trigger the subsequent recruitment of XPC–RAD23B and further downstream factors. Due to the importance of DDB2, defective cells show a remarkable reduced CPD repair activity that gives rise to the XP-E phenotype (Hwang *et al.* 1998a; Tang *et al.* 2000).

2.2 Crystal structure of human XPE with CPD damaged dsDNA

Fischer *et al.* (2011) obtained insight into the molecular basis of CPD lesion recognition by solving the crystal structure of the UV-DDB protein in complex with CPD containing DNA. For obtaining the structure, residues creating crystal contacts in the apo-protein that prevented protein crystallization were mutated. Consistent with the structure of Scrima *et al.* (2008) the CPD lesion is held by the DDB2 WD40 propeller. The crystal structure shows that the lesion recognition takes place with the help of a β -hairpin ('finger') consisting of the residues Phe371, Gln372 and His373 (Fig. 3c). This β -hairpin inserts into the minor groove of the DNA duplex at the damage site and unwinds the duplex by 12.6°. Consequently the CPD lesion is extruded into an extrahelical, flipped-out state. Finally, the DNA shows also a kink angle of about 45° at the *cis-syn* CPD lesion site.

The UV-DDB protein is part of an ubiquitin ligase complex. The displacement of COP9 signalosome (CSN) from this complex by DDB2 substrate binding leads to ubiquitination of XPC and DDB2 (Groisman *et al.* 2003; Luijsterburg *et al.* 2007). This is thought to mediate the handover to XPC by stabilizing it, resulting in an increase of its affinity for damaged DNA, while the affinity of DDB2 to DNA containing a lesion is reduced (El-Mahdy *et al.* 2006; Sugawara *et al.* 2005). Histones surrounding the lesion are ubiquitinated as well, which is suggested to slacken the nucleosome structure to promote recruitment of NER repair enzymes to the damage (Wang *et al.* 2006).

2.3 Crystal structure of human XPE with an abasic lesion site in dsDNA

Yeh *et al.* (2012) solved a dimeric crystal structure of UV-DDB (human DDB1 and DDB2) protein in complex dsDNA containing an abasic site (THF) (Fig. 3d). Each DDB2 subunit is bound to one double-stranded oligonucleotide. The DDB2 residues Asn360/Asn360 are located between the two DNA strands spanning two antiparallel β -strands (β -wing) and giving rise to a 2-fold symmetry axis of the dimer. The crystal structure shows that the residues on the leading β -strand and loop form contacts with the undamaged DNA strand and the residues on the loop and the retreating β -strand form contacts with the neighboring undamaged DNA strand. These contacts are supported by additional interactions of the N-terminal-helical domain of DDB2 (α -paddle) with the neighboring DNA-strand which lead to dimerization, explaining the high affinity to damaged DNA.

3. XPC

The human XPC protein is a polypeptide of 940 amino acids with a size of 106 kDa, forming a stable complex with the RAD23B protein (43 kDa, 409 amino acids). It is currently believed that XPC–RAD23B is responsible for the initial damage recognition with the exception of CPD lesions, where DDB2 is the first binding factor (Hey *et al.* 2002; Sugawara *et al.* 2001). The complex is known to interact with several other proteins, e.g. XPA and TFIIH (Fig. 4a, b) (Araujo *et al.* 2001; Bernardes de Jesus *et al.* 2008; Cleaver *et al.* 2009). After binding, the protein recruits further downstream factors to the damaged site (Sugawara *et al.* 1998). Biochemical assays show that the XPC–RAD23B complex binds with high affinity to different DNA lesions, e.g. 6–4PP and AAF-adducts (Batty & Wood, 2000b; Sugawara *et al.* 2002). The breakthrough came in 2001 when the Mullenders group was able to show that XPC might be the initial damage recognizer also for UV lesions. In the experiment, nuclei of living cells were irradiated with UV light at defined positions through a porous polycarbonate filter (Katsumi *et al.* 2001; Mone *et al.* 2001; Volker *et al.* 2001). The generated UV lesions could be detected by antibodies specific against CPDs and 6–4PPs. To identify which NER factors were recruited to these sites they were also visualized

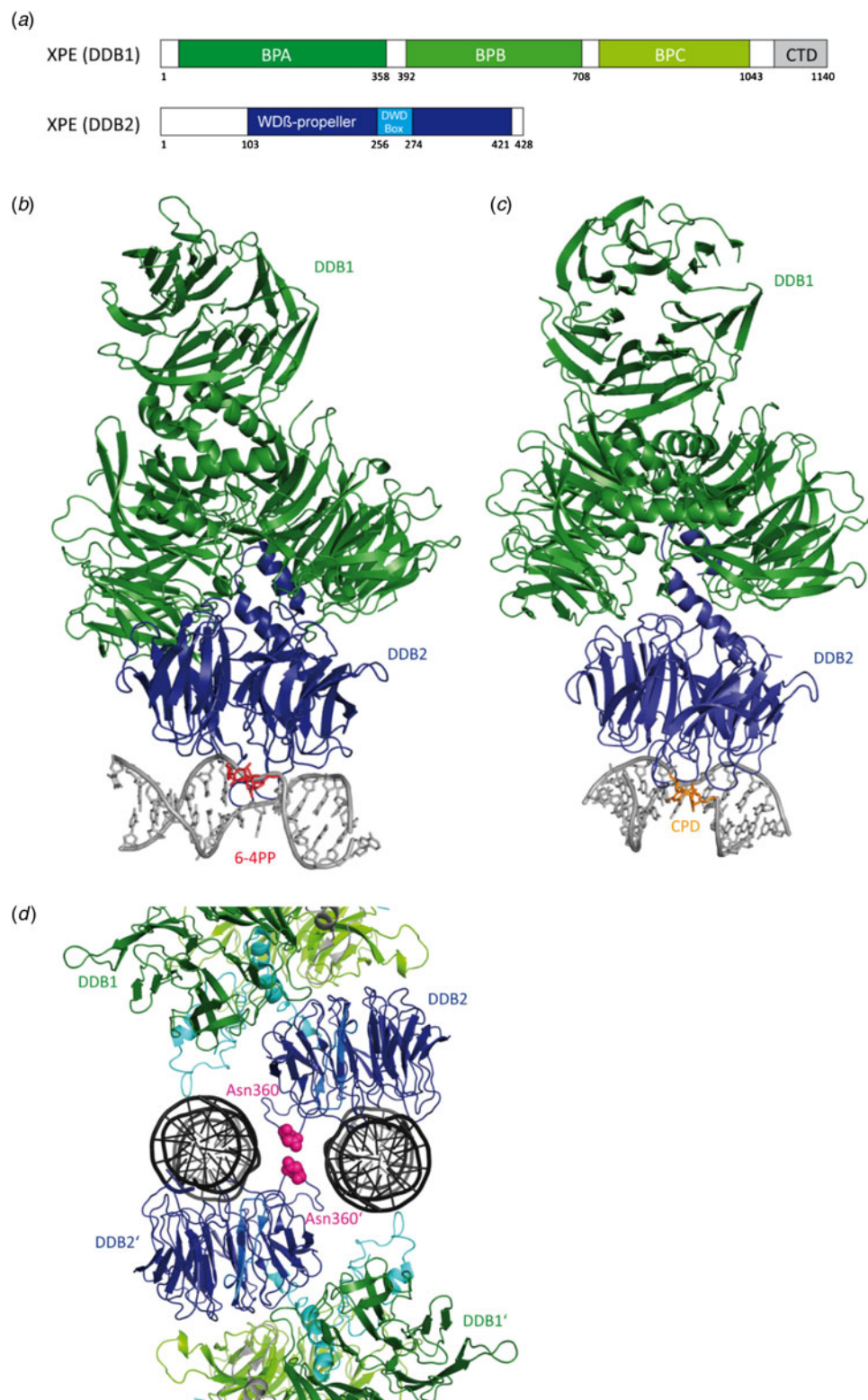


Fig. 3. Crystal structures of the human XPE protein (UV-DDB). (a) Schematic representation of the human UV-DDB protein (DDB1 and DDB2) with the different domains. (b) Crystal structure of the XPE protein in complex with 6-4PP damaged DNA (PDB code: 3EI1) (Scrima *et al.* 2008). DDB1 is depicted in green, DDB2 in blue, DNA in gray and the 6-4PP dinucleotide in red. (c) Crystal structure of XPE in complex with the CPD lesion (PDB code: 4A08) (Fischer *et al.* 2011). DDB1 is depicted in green, DDB2 in blue, DNA in gray and the CPD lesion in orange. (d) Crystal structure of the dimeric human UV-DDB in a complex with an abasic lesion (PDB code: 4e5z) (Yeh *et al.* 2012). DDB1 is depicted in green, DDB2 in blue, DNA in dark gray, DDB2 α -paddle in cyan and the amino acid Asn360 in pink.

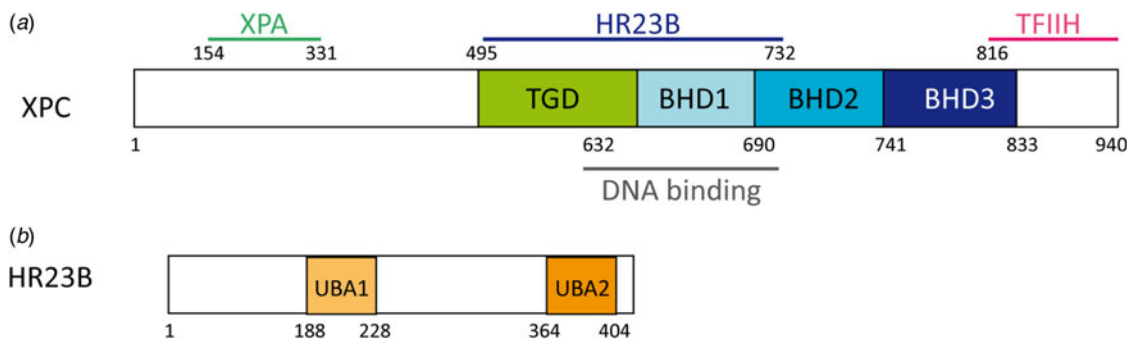


Fig. 4. Schematic representation of human XPC and RAD23B. (a) The TGD domain of XPC is depicted in green, the BHD1 domain in light blue, the BHD2 domain in blue and the BHD3 domain in dark blue. (b) The UBA1 and UBA2 domains of RAD23B are depicted in light orange and orange, respectively.

with an appropriate antibody. This experiment revealed that XPC colocalized with UV lesion sites in XPA-deficient cells, while XPA was not recruited in XPC-deficient cells (Volker *et al.* 2001). Thus, it was concluded that XPA needs the presence of XPC to be recruited to the damaged site and that XPC is the first factor being recruited to the damaged site. Further investigations pointed out that XPC is also required to unwind the DNA at the damaged site in order to recruit NER downstream factors (Katsumi *et al.* 2001; Mone *et al.* 2001). It has to be addressed that these experiments refer to the recognition of UV lesions. The general lesion recognition strategy adopted by XPC was established by a variety of experiments with different DNA lesions (Tapias *et al.* 2004). The helicase activity of XPD is not required for transcription but is indispensable for the NER process. The combined data show that XPC might be more structure specific than lesion specific. The XPC–RAD23B complex is able to recognize ssDNA (Maillard *et al.* 2007) and DNA containing destabilizing lesions. Additionally it was shown, that the complex binds the DNA lesion in an asymmetric fashion (Sugasawa *et al.* 2002; Wakasugi & Sancar, 1999). Interestingly, the complex also binds to 3–5 unpaired bases independently of the presence of damage. However, repair takes place only in the presence of DNA damage (Hess *et al.* 1997b; Sugawara *et al.* 2001).

Hence, XPC might be a sensor for unpaired base pairs that occur in bubbles and loops and binds those in complex with RAD23B (Araki *et al.* 2001; Van der Spek *et al.* 1996) and Centrin2 (Araki *et al.* 2001). This would explain why 6–4PP, AAF-dG and intrastrand cisplatin crosslinks, which introduce a severe helical kink into the DNA structure (Lukin & de Los Santos, 2006), are good substrates. On the other hand, CPDs are not recognized, possibly because they do not disturb the helical structure (Kusumoto *et al.* 2001; Reardon & Sancar, 2003; Sugawara *et al.* 2001). As mentioned previously the affinity of XPC–RAD23B to CPDs can be enhanced by introducing mismatches at the lesion site (Sugasawa *et al.* 2001).

3.1 Crystal structure of yeast XPC (Rad4) in complex with a CPD lesion

Min & Pavletich, (2007) obtained the crystal structure of the yeast homolog of XPC–RAD23B (Rad4–Rad23, Fig. 5a) in complex with a CPD lesion in a mismatched situation with two thymidines (Fig. 5b). A shorter version of the protein (the DNA-binding domain) showing the same affinity to damaged DNA as the full length protein was crystallized. The crystal structure shows that Rad4 inserts a β -hairpin (beta hairpin domain 3, BHD3)) into the destabilized DNA duplex, inducing the two damaged base pairs to completely flip out of the double helix (base flipping). A kink of 42° is created which could be confirmed by atomic force microscopy (Hwang *et al.* 1999). The protein recognizes the expelled nucleotides of the undamaged strand through Van der Waals interactions with BHD1 and transglutaminase homology domain (TGD), while the two CPD-linked nucleotides are disordered in the structure and could not be resolved. The protein undergoes a conformational change during binding to damaged DNA (*induced fit*). The authors deduced from these findings that the lesions recognized by XPC must thermodynamically destabilize the Watson–Crick double helix in a manner that facilitates the flipping-out of two bases. The protein recognizes a weakened and distorted DNA helix without binding directly to the damage.

The structure, however, does not explain why the CPD damage alone, without a mismatch, is so efficiently repaired by NER. For UV lesion repair, additional damage recognition factors must exist. New experiments show that the UV-DDB protein plays a very important role and that this protein recruits the XPC–RAD23B protein to the UV-damaged sites (Fitch *et al.* 2003; Wakasugi *et al.* 2002). Also the 6–4PP lesion which is a better NER substrate than CPD (Mitchell & Nairn, 1989), requires XPC–RAD23B as well as UV-DDB for recognition. See Section 2.

Camenisch *et al.* (2009) proposed a two-stage detection process for XPC that recognizes DNA lesions without involvement of a DNA-binding domain (BHD3). BHD1/BHD2 and a β -turn subdomain act as a damage sensor that scans the double helix

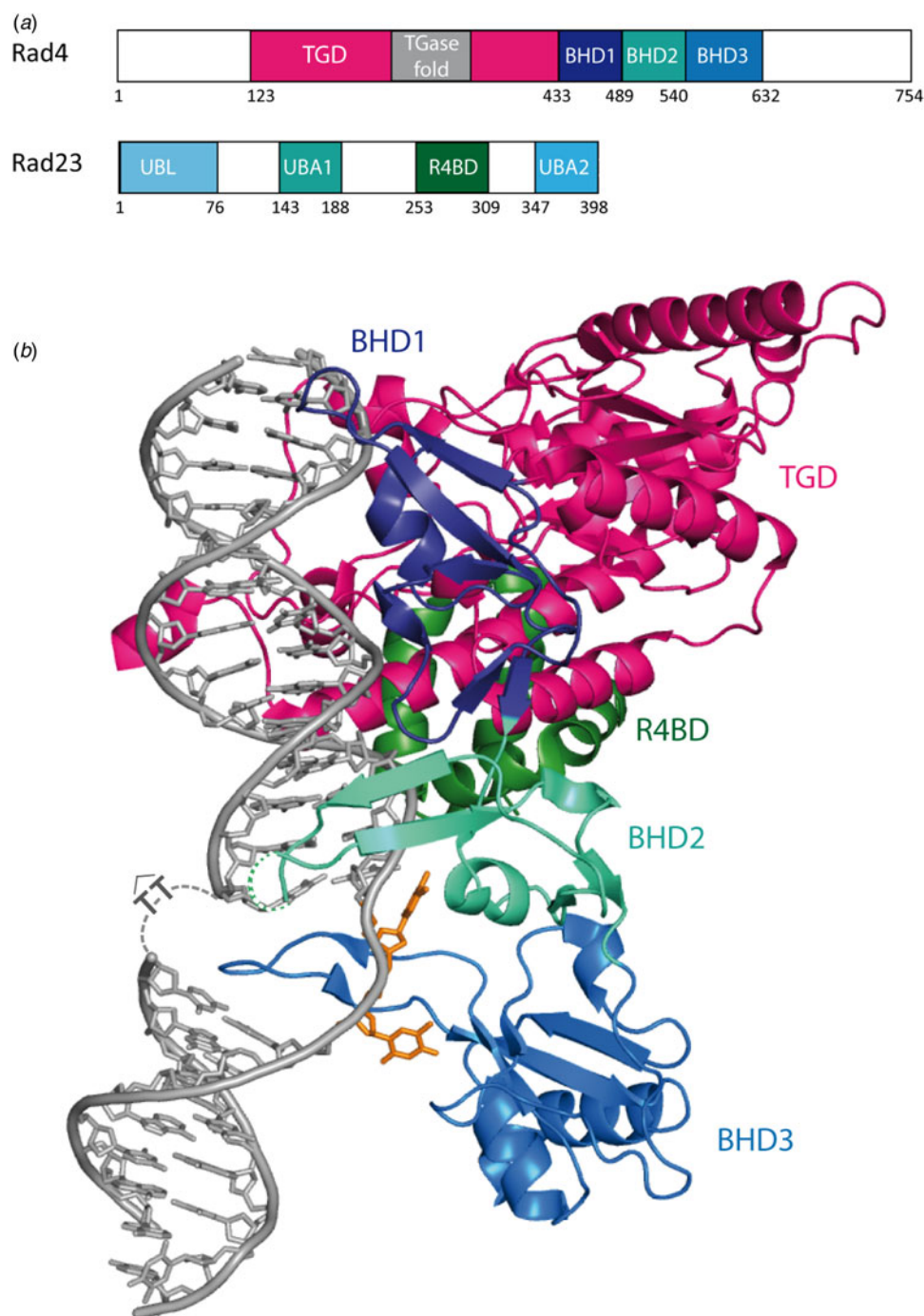


Fig. 5. (a) Schematic representation of Rad4 and Rad23. The TGD domain is depicted in pink, BHD1 in dark blue, BHD2 in cyan, BHD3 in blue and R4BD in green. (b) Crystal structure of the yeast Rad4–Rad23 in complex with a CPD damaged DNA (PDB code: 2QSG) (Min & Pavletich, 2007). Same color code as in (a). The extruded Ts in front of the CPD lesion are depicted in orange.

for unpaired bases, which forms a labile recognition complex. Subsequently, a stable recognition is formed through binding of BHD3 to the flipped out nucleotides in the undamaged strand across the lesion (Camenisch *et al.* 2009).

Chen *et al.* (2015) obtained the crystal structure of the yeast homolog of XPC–RAD23B (Rad4–Rad23) in complex with undamaged DNA using disulfide crosslinking. The crystal structure shows that Rad4 flips out normal nucleotide pairs like with damaged DNA. Therefore, structural discrimination of Rad4/XPC cannot play an important role for lesion recognition. The authors propose a ‘kinetic-gating’ mechanism for Rad4/XPC lesion recognition, which is controlled by two kinetic parameters (opening time and residence time). The weakened base stacking and hydrogen bonding within a lesion lead to a larger

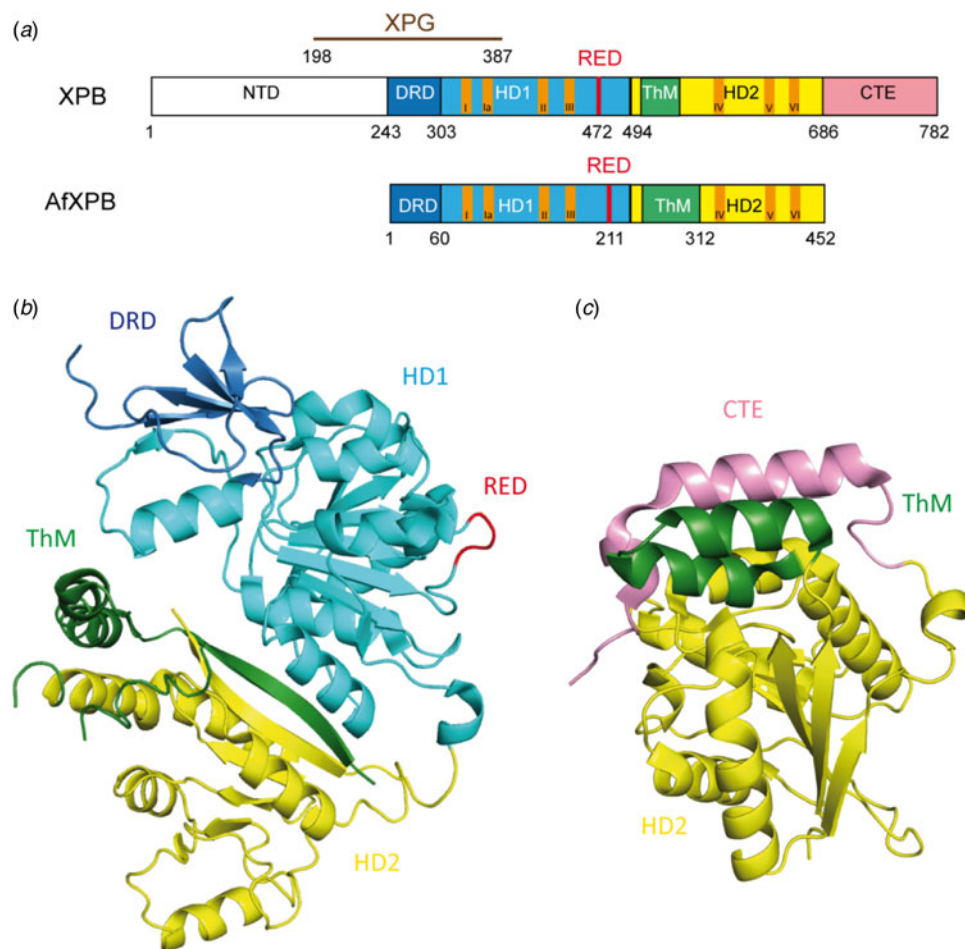


Fig. 6. (a) Schematic alignment of XPB and AfXPB showing the different domains: The DRD domain is depicted in blue, HD1 in light blue, HD1 in yellow, ThM in green and CTE (C-terminal extension) in pink. The conserved helicase motifs (Walker A motif I–VI) are depicted in orange. (b) Crystal structure of the AfXPB protein (PDB code: 2FWR). (Fan *et al.* 2006). Same color code as in (a). (c) Crystal structure of C-terminal domain of the human XPB protein (PDB code: 4ERN) (Hilario *et al.* 2013). Same color code as in (a).

opening time for damaged DNA than for undamaged DNA to form a stable recognition complex. Also, the residence time is larger for damaged DNA indicating that Rad4/XPC opens damaged DNA with a higher probability (Chen *et al.* 2015).

4. XPB

XPB (also called ERCC3) is a 3′–5′ ATP-dependent DNA helicase and a subunit of the TFIIH complex necessary for transcription and DNA repair (Gillet & Scharer, 2006; Lehmann, 2003; Schaeffer *et al.* 1993). Mutations in XPB lead to a 1000-fold increase of the melanoma risk and neurological abnormalities because of a defective TFIIH function in NER (Lehmann, 2003). The human protein consists of 782 amino acids (89 kDa). NER helicases (XPB and XPD) all include conserved Rad51/RecA like ATPase domains called helicase domains 1 and 2 (HD1 and HD2), containing the ATPase and helicase motifs characteristic for superfamily 2 (SF2) helicases (Shin *et al.* 2003; Story & Steitz, 1992). Furthermore, NER helicases consist of two accessory domains, which are responsible for transmitting the helicase domain changes upon binding to DNA. The ATPase activity of XPB is necessary during NER to anchor TFIIH to the lesion site (Abdulrahman *et al.* 2013; Coin *et al.* 1998; Dubaele *et al.* 2003; Oksenysh *et al.* 2009). Its helicase activity is, however, not needed for the repair process (Coin *et al.* 2007; Oksenysh *et al.* 2009; Tirode *et al.* 1999).

4.1 Crystal structure of *Archaeoglobus fulgidus* XPB (AfXPB)

Fan *et al.* (2006) determined the crystal structure of an *A. fulgidus* XPB homolog (AfXPB) with a sequence equivalent to residues 240–686 of the human XPB (Fig. 6b). The AfXPB structure is built up by two Rad51/RecA-like helicase domains, HD1 and HD2, containing seven conserved helicase motifs (Walker motif I, Ia and II–VI) in the middle of the polypeptide (Fig. 6a)

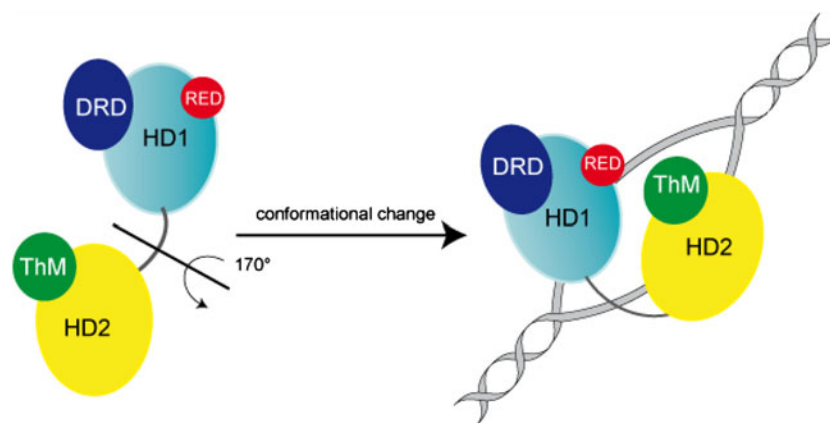


Fig. 7. Proposed model for binding of XPB to DNA (adapted from Fan *et al.* (2006)). The apo-protein is in an opened conformation. By binding to DNA, a rotation (170°) of the second helicase domain (HD2) together with the ThM domain, facilitated by HD1-mediated ATP hydrolysis, takes place and forms the closed and stable XPB–DNA complex. Same color code as in Fig. 6a.

(Fuss & Tainer, 2011; Tuteja & Tuteja, 1996). These domains use the binding and hydrolysis of ATP to induce conformational changes of the protein. Furthermore, the protein contains a DNA damage recognition domain (DRD), a unique RED motif (Arg–Glu–Asp) and a thumb domain (ThM). These domains are highly conserved and they are also found in the corresponding human protein. The structure suggests that the ATP hydrolysis provides the energy to induce a flip of about 170° in the orientation of HD2 relative to HD1 after DNA binding (Fig. 7). This conformational change brings the RED domain and the ThM domain in close vicinity and stabilizes TFIIH on the DNA by inserting a wedge (Glu473) into the double strand, gripped by the ThM domain. The motion might be important for melting the DNA duplex and anchoring TFIIH (Coin *et al.* 2007). Mutations in the RED and ThM domains as well as mutations in the ATPase domain diminish the helicase activity, suggesting that these three domains ensure the correct recruitment of TFIIH to the damaged site.

4.2 Crystal structure of human XPB

Hilario *et al.* (2013) successfully solved the crystal structure of the C-terminal fragment of human XPB (residues 494–782) containing the helicase domain 2 (HD2), the thumb-like domain (ThM) and the C-terminal extension (Fig. 6c). They investigated the structural basis for the role of the C-terminus of XPB in transcription and DNA repair. Because of lack of electron density in the structure, only the positions of the residues 502–730 could be determined. The observed structure consists of a globular domain containing a central sheet of seven parallel β -strands sandwiched by three α -helices on one side and two α -helices on the other (Fig. 6c). The C-terminal extension (residues 670–730) is built up by two α -helices linked by a loop and a β -strand, which forms the edge strand of the seven-stranded central sheet, and an unstructured tail that possibly extends to the C-terminus. The ThM domain and one α -helix of the C-terminal extension are located at the top of the central sheet. Thus, the recruitment of TFIIH to the damage site is an active process controlled by the ATPase motif of XPB. This suggests that this subunit acts as an ATP-dependent hook that stabilizes the binding of the TFIIH to damaged DNA (Hilario *et al.* 2013). In fact, a mutated ATPase activity of XPB abolishes the NER activity (Guzder *et al.* 1994; Sung *et al.* 1988). Recruiting TFIIH to the damage site by the ATPase activity of XPB may reorganize the protein–DNA complexes in order to allow new protein–protein or protein–DNA contacts to happen.

5. XPD

The XPD (also called ERCC2) protein with a size of 87 kDa (760 amino acids) is known to be involved in transcription and DNA damage repair. It is a key subunit of the TFIIH complex and acts as an ATP-dependent 5′–3′ DNA helicase that belongs to the superhelicase 2 family (SF2) proteins (Araujo *et al.* 2000; Coin *et al.* 2007). The helicase activity of XPD is not required for transcription but is indispensable for the NER process (Winkler *et al.* 2000). The TFIIH complex interacts with other NER factors such as XPA, XPC, XPG and XPF–ERCC1 (Evans *et al.* 1997a; Gillet *et al.* 2005). Mutations in human XPD lead to three different inherited diseases: XP, the CS and TTD. XPD is thought to be responsible for the damage verification (Mathieu *et al.* 2010; Naegeli *et al.* 1993) before the recruitment of downstream NER factors to the lesion site occurs. This idea has been postulated based on the observation that the helicase activity of XPD becomes inhibited by the presence of DNA damage (Mathieu *et al.* 2010, 2013; Naegeli *et al.* 1992).

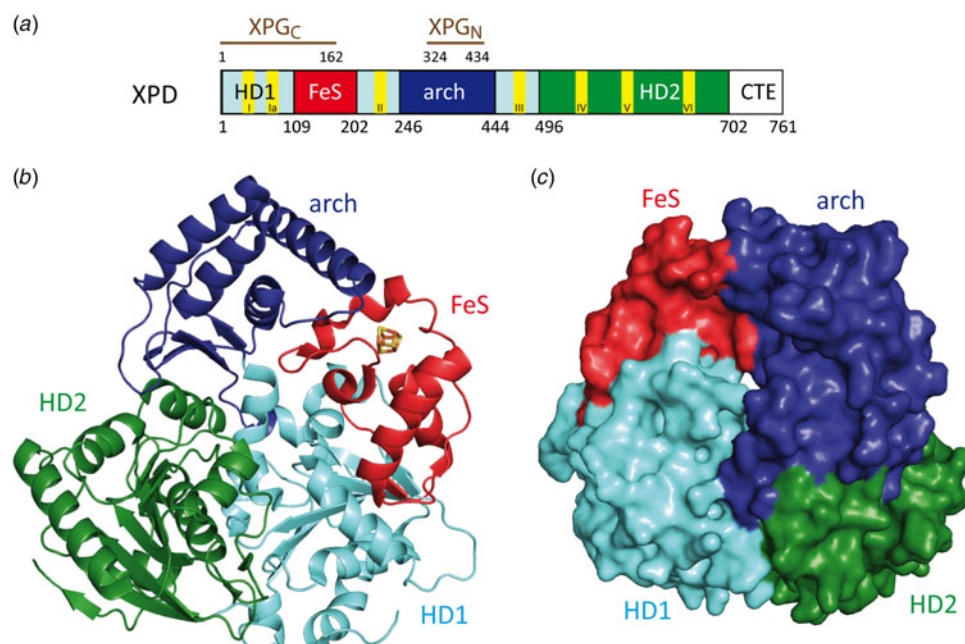


Fig. 8. (a) Schematic representation of human XPD (Fan *et al.* 2008). The protein interacts with XPG_C (C-terminal domain of XPG) and XPG_N (N-terminal domain of XPG) depicted in brown. (b) Crystal structure of SaXPD (PDB code: 3CRV) (Fan *et al.* 2008) with the HD1 domain (light blue) and the HD2 domain (green) form the ATP-binding interface. The arch domain is depicted in blue and FeS domain in red. The iron-sulfur cluster (4Fe4S) is depicted in orange and yellow. (c) Surface representation of SaXPD showing the tunnel formed by the different domains, which is large enough to pass through ssDNA.

5.1 Crystal structure of *Sulfolobus acidocaldarius* XPD (SaXPD)

Several crystal structures of the catalytic domain of XPD from different organisms (*S. acidocaldarius* (Fan *et al.* 2008) (SaXPD; Fig. 8), *Thermoplasma acidophilum* (Wolski *et al.* 2008) (TaXPD) and *Sulfolobus tokodaii* (Liu *et al.* 2008)) reveal a four domain organization: a domain harboring the 4Fe4S iron-sulfur cluster (FeS domain), an arch domain and two canonical Rad51/RecA like domains, the helicase domains HD1 and HD2 (Fig. 8a). The FeS domain is essential for the helicase activity of XPD.

The different domains form a tunnel that is large enough to harbor ssDNA (Fig. 8b, c) (Fan *et al.* 2008; Liu *et al.* 2008; Wolski *et al.* 2008). Fan *et al.* proposed a model of XPD binding to ssDNA (Fig. 9) (Fan *et al.* 2008). The iron-sulfur cluster is supposed to act as a molecular plough, which scans blocking damages by breaking hydrogen bonds (Lehmann, 2008; White, 2009). In this process, one strand of the duplex is supposed to be pulled through the internal channel with the help of the ATP-dependent motor domain, while damaged DNA will plug the channel (Fan *et al.* 2008; Liu *et al.* 2008; Mathieu *et al.* 2013; Pugh *et al.* 2012; Wolski *et al.* 2008). In this way, the protein undergoes a conformational change by closing the domain between the arch and the FeS motifs. The helicase domains HD1 and HD2 furnish the necessary flexibility for translocation (Fan *et al.* 2008).

Because the channel is not narrow enough to distinguish between damaged and undamaged bases 2008, Wolski *et al.* proposed that during the unwinding process single bases are analyzed through insertion of the DNA backbone into a neighboring binding pocket on the protein surface. Damaged bases would not fit into the pocket, which could be the signal that leads to blocking of the helicase activity (Wolski *et al.* 2008).

5.2 Crystal structure of *T. acidophilum* (TaXPD) in complex with ssDNA

Finally, in 2012 Kuper *et al.* successfully solved the crystal structure of TaXPD in complex with a ssDNA oligonucleotide (22mer), shedding light on the mechanism of polarity and translocation of the protein along the DNA (Kuper *et al.* 2012). At the same time Pugh *et al.* showed that the 5'-3' translocation polarity needs conserved residues in the HD1 and the FeS domains. They propose that the FeS domain provides a wedge-like feature enabling the duplex opening during unwinding (Pugh *et al.* 2012).

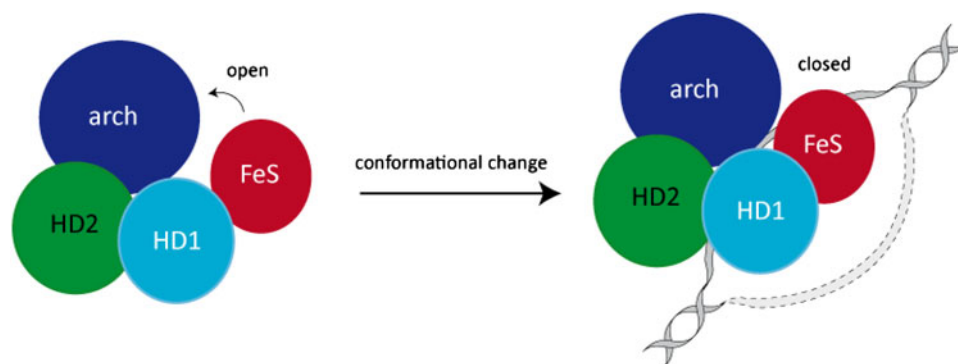


Fig. 9. DNA-binding model of SaXPD adapted from Fan *et al.* (2008). The HD1 domain is depicted in light blue, the HD2 domain in green, the arch domain is depicted in blue, the FeS domain in red and the DNA in gray. Upon binding to the DNA the protein undergoes a conformational change by closing the domain between the arch and the FeS domains.

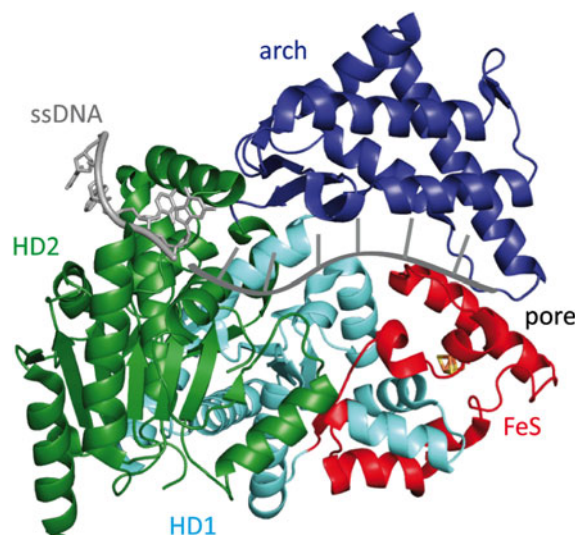


Fig. 10. Crystal structure of *Thermoplasma acidophilum* XPD (TaXPD) in complex with a 10 nt ssDNA. The arch domain is depicted in dark blue, the FeS domain in red, the HD1 domain in cyan and the HD2 domain in green. The four bases (TACG), showing electron density, are depicted as gray sticks. The modeled ssDNA in the pore is schematically represented in gray. The modeled structure is adapted from Kuper *et al.* (2012).

The structure allowed the Kisker group to define the translocation path of the DNA substrate through the protein and to determine the DNA-binding domain (DBD) of TaXPD by mutational and biochemical analysis. Only four bases (TACG) of the oligonucleotide could be identified in the electron density. These bases are located in the cleft of the HD2 domain (Fig. 10). This domain was already predicted as a potential binding cleft by Wolski *et al.* (2008). The structure shows stacking of Trp549 with the adenine of TACG. Arg584 builds hydrogen bonds with N1 and N2 of the phosphate backbone of adenine and cytosine. Phe538 does not directly contact the DNA, but it is supposed to stabilize larger bases like adenine or guanine. A fourth residue Asp582 stabilizes Arg584 by hydrogen bonding to maintain it in position for DNA binding. The 5' end points away from the HD2 domain and the 3' end points towards HD1 domain. Kuper *et al.* propose a model (Fig. 10) which shows how the ssDNA is situated and translocated in the protein.

5.3 Proposed model of damage verification

Based on these new data about XPD, Fuss and Tainer proposed an extended model of the damage verification process, the so-called 'bind-pry-unwind'-model: (Fuss & Tainer, 2011) here XPB, XPD and CAK of the TFIIH complex are all considered to be essential for the coordination of damage recognition, verification, repair and signaling. In this model, the damage recognition proteins not only recruit the TFIIH complex and other NER factors to the damage site, but also fix the DNA at the same time. Kim *et al.* suggested that XPB facilitates the unwinding of the DNA duplex during transcription by distorting the



DNA underneath the transcription bubble, which is fixed by the transcription factors and RNAP II (Kim *et al.* 2000). Similarly XPB could unwind and open the DNA. In that way, the damage verification step and the DNA unwinding could be linked. These findings are in accord with the low helicase activity of XPB and its strong conformational change upon binding to DNA (Fan *et al.* 2006). Furthermore, TFIIH can thereby position itself directly next to XPC–RAD23B or RNAP II, which establishes a common mechanism of how XPB may act in GG-NER, TC-NER and transcription. Through its short binding canal, XPB could place TFIIH vertically to the DNA and 5' of the recognition protein. Then XPB could break the DNA strand, so that TFIIH bends it towards the DNA enabling XPD to bind about 22 base pairs away from the lesion on the damaged strand. Fuss and Tainer assume that the distance between XPB and XPD determines the size of the open DNA-bubble, corresponding to about 27 base pairs that are eliminated during NER. XPA catalyzes the release of CAK from the TFIIH complex (Coin *et al.* 2008). This step signals that the repair process can be continued and promotes the incision step. The XPD helicase unwinds the duplex until the 4Fe4S-cluster docks onto the damage and is blocked by it (Mathieu *et al.* 2010; Naegeli *et al.* 1993). Recent studies show that XPD exists in different conformations depending on the type of damage (Buechner *et al.* 2014). After the blockage, XPD and XPC–RAD23B anchor the TFIIH complex and build a platform for binding of ERCC1–XPF and XPG (Araujo *et al.* 2001).

Following this model RPA, XPA and XPG probably reach the lesion site independently and build together with TFIIH the pre-incision complex (Rademakers *et al.* 2003; Riedl *et al.* 2003; Volker *et al.* 2001; Wakasugi & Sancar, 1998). Here, all NER factors, except ERCC1–XPF, are present and stably bound to the open DNA bubble.

Recently, the Kisker group showed by mutational analysis that the FeS domain of XPD is crucial for NER (Kuper *et al.* 2014). Furthermore, the successful repair process is dependent on the enzymatic activities of XPD which involves DNA binding, ATPase activity, helicase activity and the correct interaction with other TFIIH subunits. In 2013, the Naively group showed by uncoupling the DNA unwinding activity from the damage verification activity of XPD, that residues Tyr192 and Arg196 form a sensor pocket located ahead of the narrow central protein pore responsible for damage recognition. By mutating these two residues, the helicase and ATPase activities are not affected but the damage sensing during ATP-driven scanning movement along nucleic acid lattices fails and the ability of XPD to form stable DNA damage recognition intermediates is significantly reduced (Mathieu *et al.* 2013). The activity of XPD however is not relevant for transcription and DNA binding.

6. XPA

The XPA protein (xeroderma pigmentosum group complementation group A) consists of 273 amino acids (31 kDa) and can occur as either a homodimer or in complex with other NER proteins (Liu *et al.* 2005; Yang *et al.* 2002). The protein is constituted of several domains (Camenisch *et al.* 2006; Camenisch & Nägeli, 2008): the N-terminal domain (XPA_{4–97}) includes the RPA34- and ERCC1-binding regions (He *et al.* 1995; Li *et al.* 1995a). The C-terminal domain (XPA_{226–273}) interacts with TFIIH (Camenisch & Nägeli, 2008; Ikegami *et al.* 1998; Park *et al.* 1995a). The central domain (XPA_{98–219}) is the minimal DBD (XPA-DBD) (Cleaver & States, 1997; Kuraoka *et al.* 1996). The protein contains a zinc-finger binding domain (Cys105–X₂–Cys108–X₁₇–Cys126–X₂–Cys129) (Hess *et al.* 1998). The C₄–zinc finger sustains the stability of the protein and plays a role in the interaction with RPA70 (Fig. 11a) (Ikegami *et al.* 1998; Kuraoka *et al.* 1996; Morikawa & Shirakawa, 2000; Morita *et al.* 1996; Sugitani *et al.* 2014). Mutation of one or all zinc-binding cysteines leads to a loss of function of XPA and to an almost complete loss of function of NER (Li *et al.* 1995b; Mellon *et al.* 1987).

It has been demonstrated that XPA binds ssDNA, dsDNA and denaturated stretches of DNA in cooperation with RPA (Buschta-Hedayat *et al.* 1999; Missura *et al.* 2001; Wang *et al.* 2000; Yang *et al.* 2006). XPA shows binding affinity to irradiated DNA (Jones & Wood, 1993; Liu *et al.* 2005), DNA containing bulky adducts and cisplatin lesions (Jones & Wood, 1993; Liu *et al.* 2005), as well as to DNA duplexes containing loops and bulges (Camenisch *et al.* 2006). Camenisch *et al.* could show by mutational analyses that the residues Lys141 and Lys179 found on the surface of XPA next to the DNA-binding pocket are responsible for the sensing of DNA deformability as the common characteristics of damaged sites (Camenisch *et al.* 2006; Miyamoto *et al.* 1992). It was postulated that these two lysines interact with the DNA backbone, when XPA is in complex with RPA and kinked DNA (Read *et al.* 1993; Weir *et al.* 1993). Photocrosslinking experiments demonstrated that the mutated protein does interact with kinked DNA structures (Camenisch *et al.* 2006). This led to the assumption that the protein scans the DNA duplex for the presence of the kinked DNA structure with its positively charged pocket. The loop-rich domain of XPA was determined to be responsible for DNA binding. The ternary structure is stabilized by the zinc-finger domain (Miyamoto *et al.* 1992; Yan *et al.* 2002).

6.1 Nuclear magnetic resonance (NMR) structure of the DBD of human XPA

Ikegami *et al.* (1998) determined the NMR structure of the DBD of XPA (XPA_{98–219}, called XPA-DBD) and they found that the protein contains α -helices and β -sheets (Fig. 11b). The structure shows the presence of a positively charged (basic amino

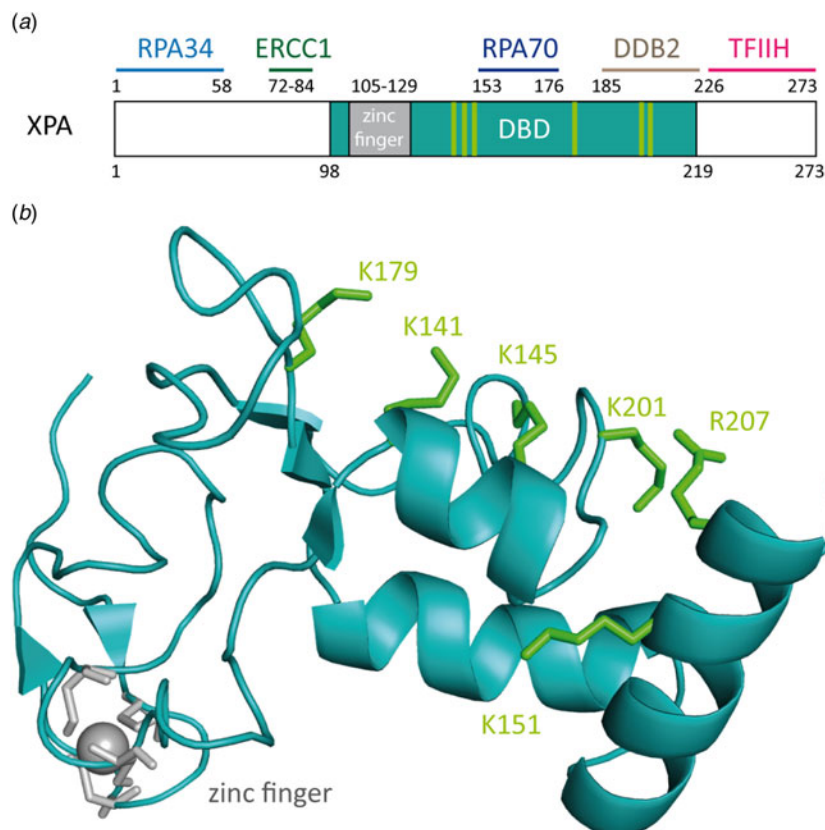


Fig. 11. (a) Schematic view of the human XPA protein with XPA-DBD (DNA-binding domain, XPA₉₈₋₂₁₉) in cyan the zinc-finger domain in gray. The binding domains of XPA with other NER proteins are depicted in light blue for RPA34, green for ERCC1, dark blue for RPA70 and pink for TFIIH. (b) NMR-structure of the XPA-DBD (Buchko *et al.* 2001; Camenisch *et al.* 2006; Ikegami *et al.* 1998) (PDB code: 1XPA). Positively charged amino acids that build the cleft are depicted as green sticks. The zinc atom is shown in gray.

acids) cleft with the right dimension to bind dsDNA (Buchko *et al.* 1998, 2001). This positively charged cleft was speculated to interact with the negatively charged DNA backbone.

Sugitani *et al.* (2014) redefined the DBD of XPA. By extending the originally characterized XPD-DBD at its C-terminus by 20 amino acids (XPA₉₈₋₂₃₉), they improved the DNA-binding properties. The new construct with the additional critical basic residues and helical elements at its C-terminus was found to bind to DNA substrates with the same affinity as the full-length protein. The group reported, based on NMR data, that XPA₉₈₋₂₃₉ contains the same globular core as XPA₉₈₋₂₁₉ and similarly undergoes a conformational change when binding to Y-shaped ss/ds-DNA junctions.

The rather small XPA protein interacts with a large number of other NER factors which supports its function as a scaffold protein for the assembly of the incision complex during repair (Gillet & Scharer, 2006; Naegeli & Sugasawa, 2011). Probably, XPA binds best to already prekinked DNA structures, which explains why XPA is able to bind defect Watson-Crick base pairing, created by mismatches, bubbles or three-way junctions (Bootsma & Hoeijmakers, 1993; Buschta-Hedayat *et al.* 1999; Missura *et al.* 2001). These structures can form easily kinked DNA structures. XPA shows a slight binding preference to UV-damaged DNA and bulky adducts over undamaged DNA (Jones & Wood, 1993; Robins *et al.* 1991; Thoma & Vasquez, 2003). Similarly to XPC, XPA recognizes 6-4PP lesions but not CPDs (Jones & Wood, 1993). CPD might block the bending of major grooves because of its stiff four-membered ring which points into the major groove. XPA shows a high affinity to DNA containing aromatic adducts and cisplatin crosslinks. Cisplatin crosslinks also kink the DNA structure. The damage recognition capability of XPA increases with RPA (Jones & Wood, 1993; Li *et al.* 1995a; Mellon *et al.* 1987) and ERCC1 (Nocentini *et al.* 1997; Saijo *et al.* 1996). XPA does not bind ssDNA, in agreement with the observed size of the positively charged cleft able to fit dsDNA (Buchko *et al.* 1998; Ikegami *et al.* 1998). For single-base bulky adduct lesions such as AAF-dG, it is unclear to which extent they prekink the duplex. Bulky adducts can increase the local conformational flexibility of DNA (Isaacs & Spielmann, 2004). XPA might employ a DNA flexibility scanning mechanism to create a kink, which could be stabilized by intercalation of the aromatic ring into the duplex and may be the secret behind the broad substrate

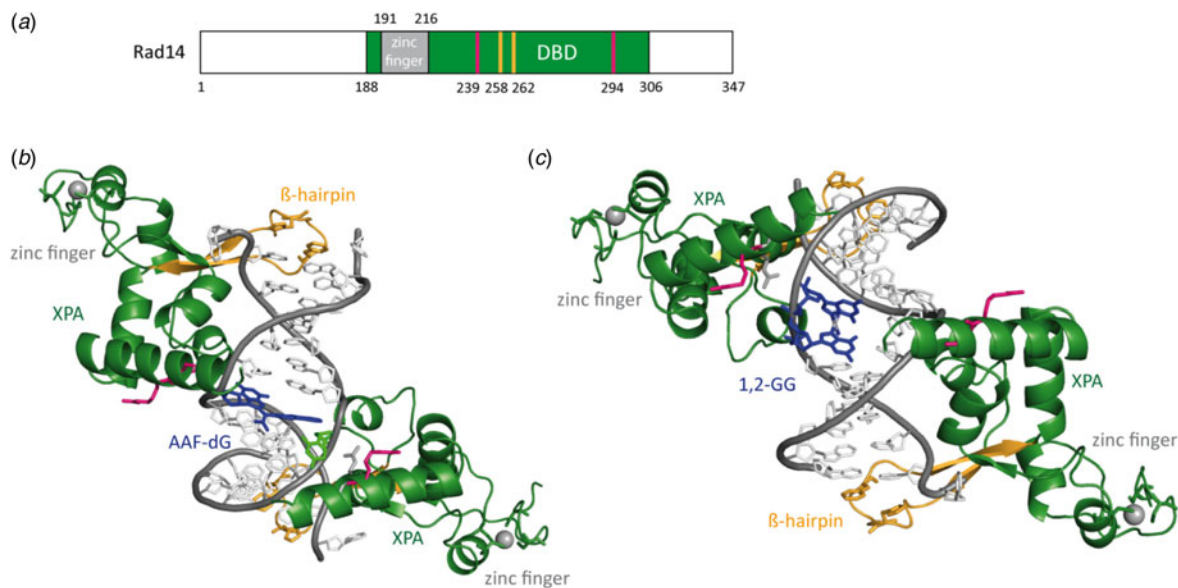


Fig. 12. (a) Schematic sequence of Rad14 with the DBD depicted in green, the zinc-finger domain in gray and the residues interacting with the DNA in orange and pink. (b) Crystal structure of Rad14 in complex with AAF-dG crosslinked DNA (PDB code: 5A3D) (Koch *et al.* 2015). Same color code as in (a). The AAF-dG residue is depicted in dark blue. (c) Crystal structure of Rad14 in complex with 1,2-GG cisplatin (PDB code: 5A39) (Koch *et al.* 2015). Same color code as in (a). The 1,2-GG cisplatin crosslink is depicted in dark blue.

promiscuity of NER. Koch *et al.* could obtain two crystal structures of Rad14 with a bulky lesion, which confirm this hypothesis (Koch *et al.* 2015).

6.2 Crystal structures of Rad14 with AAF-dG and cisplatin containing dsDNA

Recently, Koch *et al.* (2015) solved two crystal structures of the yeast homolog of XPA (Rad14) in complex with AAF-dG and 1,2-GG cisplatin containing DNA (Fig. 12b, c). The DBD of the protein (residues 188–306) was crystallized with damaged DNA (Fig. 12a). The structures allow us to uncover the molecular mechanism of XPA interacting with damaged DNA and are consistent with previous biochemical data showing the binding of XPA to kinked DNA structures (Buschta-Hedayat *et al.* 1999; Camenisch *et al.* 2006). The protein assumes an overall α/β -fold quite similar to its human homolog (Ikegami *et al.* 1998; Koch *et al.* 2015). Comparable with Rad4, Rad14 does not bind the lesion directly; instead it recognizes a weakened DNA duplex structure (Min & Pavletich, 2007). By the binding of two Rad14 proteins to the duplex, the DNA undergoes a sharp kink of 70° at the lesion site and single stranded regions are formed above and below the 13mer binding motif. Through insertion of a β -hairpin (His258 and Phe262) from each protein (acting like ‘fingers’) exactly six base pairs away from the lesion, a 13mer recognition motif is generated. Another domain composed of helix $\alpha 7$ (‘thumb’ containing Arg234 and Arg239) interacts with the DNA backbone close to the lesion, so that DNA bending takes place into the major groove. The lesion is located inside the 13mer duplex instead of being flipped out. Because of the relatively few interactions between Rad14 and the duplex, the protein might preferably bind already pre-kinked DNA structures (Camenisch *et al.* 2006). The 2:1 complex can coordinate contacts between TFIIH, DNA and RPA and recruit ERCC1–XPF (Li *et al.* 1995a; Park *et al.* 1995b; Tripsianes *et al.* 2007).

7. XPF–ERCC1

The human DNA repair endonuclease XPF (gene called ERCC4) is a protein of 916 amino acids (104 kDa) (Brookman *et al.* 1996). It forms a stable ERCC1–XPF heterodimer enzyme complex that participates in DNA repair and DNA recombination (Sijbers *et al.* 1996; Tsodikov *et al.* 2005). ERCC1 is a protein of 297 amino acids (32.5 kDa) that mediates DNA–protein and protein–protein interactions. The XPF–ERCC1 complex cuts specific structures near an ss/dsDNA junction at the 5′-end of single stranded damaged DNA strand and hence makes the first incision in the NER process. Its nuclease activity is absolutely essential in the NER process (de Laat *et al.* 1999; Sijbers *et al.* 1996). In the complex, XPF provides the endonuclease active site and is involved in DNA binding and establishing additional protein–protein interactions. XPF–ERCC1 is known to be the last

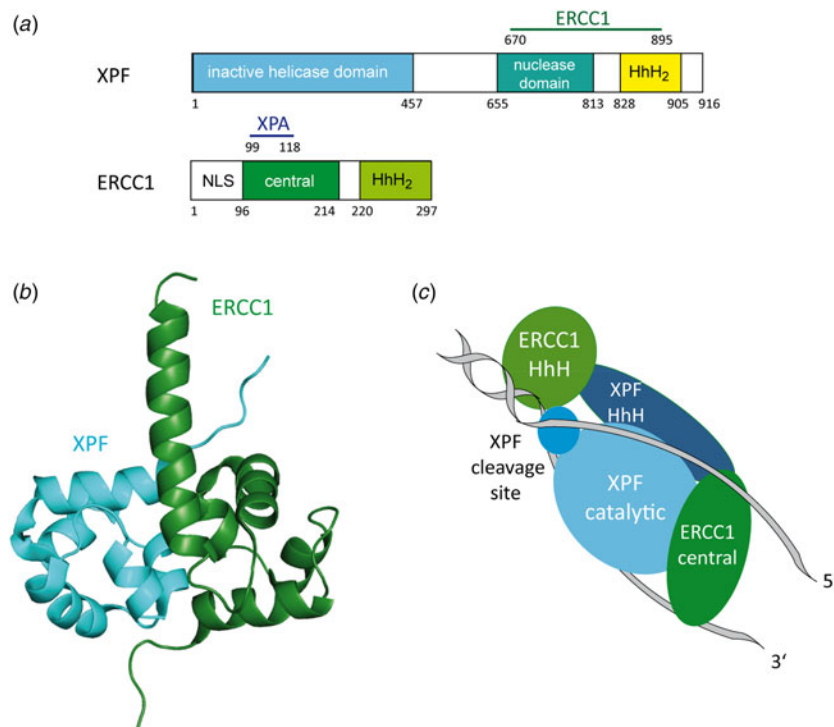


Fig. 13. (a) Structural domains of the human XPF and ERCC1 proteins. XPF interacts with ERCC1 (residues 670–895) depicted in green. The inactive helicase domain is depicted in light blue, the nuclease domain in cyan and the HhH₂ domain in yellow. ERCC1 interacts with XPA (residues 99–118) depicted in dark blue. The central domain is depicted in green and the HhH₂ domain in light green. (b) Structure of the human XPF–ERCC1 heterodimer (PDB code: 2A1J) (Tsodikov *et al.* 2005) with ERCC1 depicted in green and XPF depicted in light blue. (c) Model of the XPF–ERCC1 complex binding to a splayed DNA substrate (depicted in gray). The ERCC1 central domain and HhH domain are depicted in green. XPF, XPF cleavage site and HhH domain are depicted in light blue.

factor joining the NER complex before dual incision takes place (Riedl *et al.* 2003; Volker *et al.* 2001; Wakasugi & Sancar, 1998). Evidence exists that it is recruited by XPA (Fig. 13a) (Li *et al.* 1994, 1995a; Park & Sancar, 1994; Saijo *et al.* 1996).

7.1 Crystal structure of human XPF–ERCC1

Tsodikov *et al.* (2005) solved the crystal structure of the minimal dimerization domains of the human XPF and ERCC1 (Fig. 13b). The crystallized domain shows a similar endonuclease activity compared to that of the full-length complex. Both protein domains exhibit double helix–hairpin–helix motifs (HhH₂) related by a pseudo-2-fold symmetry axis (Tripsianes *et al.* 2005). The HhH₂ motifs are known to be ds- as well as ss-DNA-binding motifs (Newman *et al.* 2005; Singh *et al.* 2002). Both C-terminal HhH₂ domains of XPF and ERCC1 are essential and sufficient for a stable heterodimer formation (de Laat *et al.* 1998). XPF might play a role as a scaffold protein, allowing proper folding of the ERCC1 domain. All XPF family members contain a catalytic domain comprising the active site motif (GDX_nERKX₃D) and they need a divalent cation for their nuclease activity (Nishino *et al.* 2003; Sijbers *et al.* 1996).

7.2 NMR structure of C-terminal human XPF in complex with ssDNA

Das *et al.* (2012) solved the NMR structure of the C-terminal domain of the human XPF protein (residues 823–905) in complex with a 10-nucleotide ssDNA fragment. In earlier NMR studies of the apo-protein, the group showed that the XPF homodimer complex is more stable than the XPF–ERCC1 heterodimer complex (Das *et al.* 2008). The structure shows that a positively charged region of one HhH domain interacts with the ssDNA phosphate backbone. The central guanine (G5) is extruded from the duplex and inserted into a pocket formed by residues from both HhH motifs of XPF (Das *et al.* 2008; Tripsianes *et al.* 2005). This cavity offers a large interaction surface area for the protein with the DNA. The DNA forms a right-handed structure binding to the protein, which mainly interacts with the ssDNA phosphate backbone (Fig. 14). In addition, the protein interacts with three bases: thymidine3 (His857), thymidine4 (Ser875, Ile876) and G5 (Leu877, Gly878 and Asn879). Combined with previous data (Tripsianes *et al.* 2005) the group proposes a binding model of XPF–ERCC1 to ss/dsDNA junctions (Das *et al.* 2012). They modeled the XPF–ERCC1 complex bound to a splayed-arm DNA substrate showing

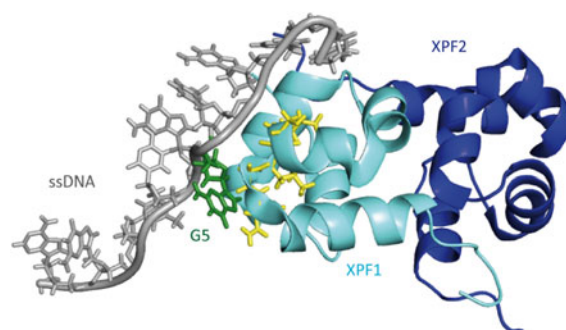


Fig. 14. NMR structure of the C-terminal domain of human XPF in complex with an ssDNA fragment (PDB code: 2KN7) (Das *et al.* 2012). The first XPF (XPF1) protein is depicted in light blue showing the interacting residues with the DNA in yellow. The second XPF (XPF2) protein is depicted in light blue. The central guanine (G5) is depicted in green.

that the conserved HhH domain of ERCC1 contacts the phosphate backbone of dsDNA in the minor groove with its two hairpins H1 and H2 while the XPF HhH domain stabilizes the ssDNA via interactions with a 5'/3' polarity (Fig. 13c). The damaged strand gets cleaved by the catalytic residues of the XPF nuclease domain and the ERCC1 central domain connects with XPA and ssDNA on the NER complex (de Laat *et al.* 1998).

8. XPG

XPG (gene called ERCC5) is a single-stranded structure-specific DNA 3'-endonuclease involved in NER. It belongs to the family of FEN-1 endonucleases (Flap endonuclease) (Lieber, 1997). The human protein consists of 1186 amino acids (133 kDa) (O'Donovan *et al.* 1994b) and it interacts with several other NER proteins (Fig. 15). Mutations in this gene cause the XP-G phenotype, characterized by hypersensitivity to UV light and increased susceptibility for skin cancer development following UV exposure. Furthermore, it can lead to the CS characterized by severe growth defects and mental retardation (Constantinou *et al.* 1999). The XPG protein cleaves DNA near junctions of dsDNA to ssDNA with a defined polarity moving 3'-5' away from the junction. It has been observed that XPG needs a duplex opening of more than five base pairs for efficient cleavage (Evans *et al.* 1997a). The protein is known to interact with XPB, XPD, RPA and PCNA (Feldes & Bonatto, 2015; Iyer *et al.* 1996). XPG is needed for the proper recruitment of XPF-ERCC1 to the preincision complex. Once the 5'-incision is made by the XPF-ERCC1 nuclease, XPG becomes fully active (Staresincic *et al.* 2009). Despite this knowledge, the exact mechanism of substrate recognition by XPG is still unclear.

The human protein possesses a globular structure like the FEN-1 endonuclease (Hosfield *et al.* 1998; Hwang *et al.* 1998b). It contains an N-terminal domain and an I-region linked by a 600 residue long spacer region (Dunand-Sauthier *et al.* 2005; Hohl *et al.* 2007). The N-terminal and the I-region are evolutionary conserved among similar endonucleases (T4 RNaseH, Taq Polymerase). The spacer region is not involved in catalysis. It is responsible for proper DNA cleavage and for the recruitment and positioning of XPG to the lesion site (Constantinou *et al.* 1999; Dunand-Sauthier *et al.* 2005; He *et al.* 1995; Thorel *et al.* 2004). It is also known to interact with RPA and TFIIH. The protein contains a D1 box (residues 184–210) and a D2 box (residues 890–984) which are evolutionary conserved (Dunand-Sauthier *et al.* 2005; Houle & Friedberg, 1999). One cation is required for the endonuclease activity in the active site. The protein shows a higher affinity for Mn^{2+} than for Mg^{2+} and uses the more abundant one for catalysis (Hohl *et al.* 2007; Iyer *et al.* 1996). Two residues in the active site (Glu791 and Asp812) coordinate the divalent cations (Constantinou *et al.* 1999) (light blue sticks in Fig. 15). The protein incises flap DNA and bubbled DNA structures with a defined polarity (Evans *et al.* 1997b; Hohl *et al.* 2003; O'Donovan *et al.* 1994a; Tsutakawa *et al.* 2011).

8.1 Crystal structure of the yeast XPG (Rad2) in complex with dsDNA

To understand the mechanism of DNA binding by XPG Mietus *et al.* (2014) solved in 2014 the crystal structure of the catalytic core of the *Saccharomyces cerevisiae* homolog of XPG (Rad2- Δ SC, Fig. 16a) in complex with four different DNA substrates. For crystallization a XPG variant missing the mostly disordered spacer residues 111–730 between the two sequence segments N and I of the nuclease domain was used. The double-stranded oligonucleotides (12–15 bp) were designed with single-stranded overhangs constructed with one to six adenines or thymidines. In the crystal structure two Rad2 proteins bind to one DNA substrate with a single-protein molecule covering 13 bp of the double-stranded portion. Rad2 utilizes



Fig. 15. Schematic representation of human XPG with its different domains and the region interacting with other NER proteins. The D1 box is depicted in light green, the D2 box in green and the 600 aa spacer region in orange. The XPB and XPD interacting region is highlighted in purple, RPA interacting region in dark blue and the PCNA interacting region in pink. Glu791 and Asp812 in the active sites are depicted as light blue bars.

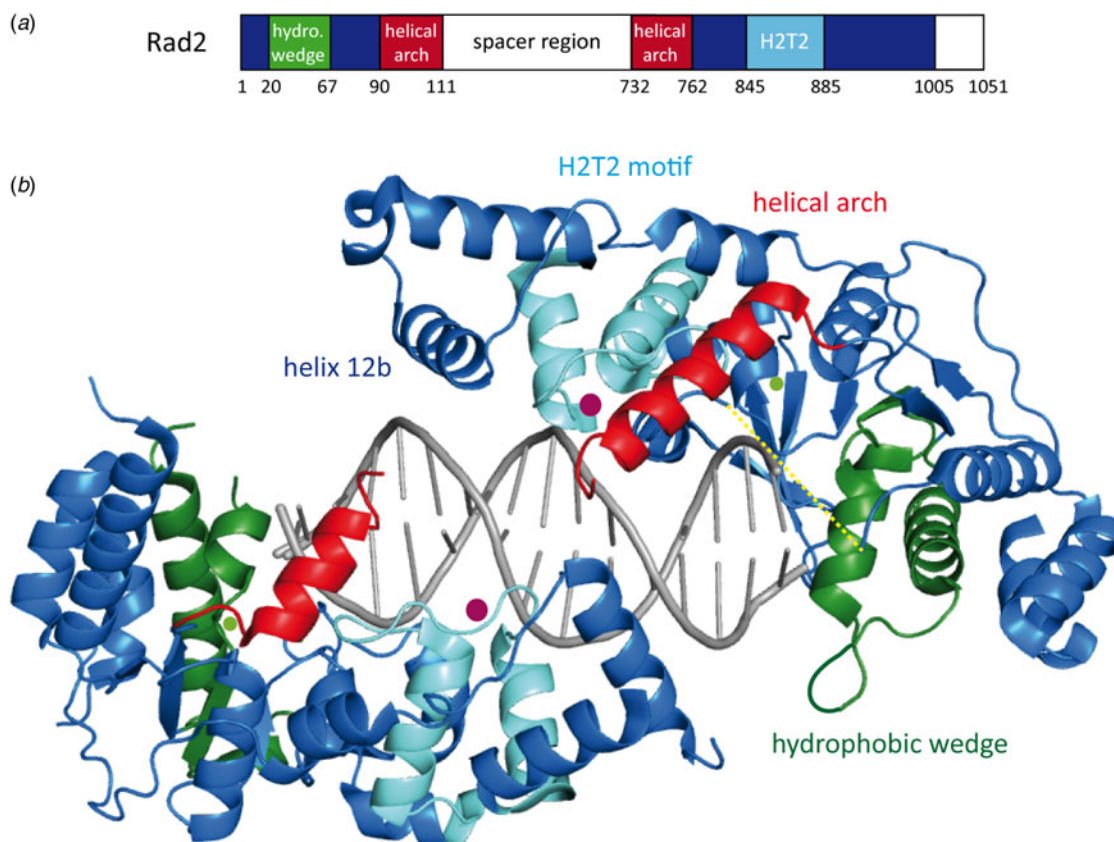


Fig. 16. Crystal structure of the Rad2–DNA complex (PDB code: 4Q0W) (Mietus *et al.* 2014). (a) Schematic representation of the protein. The hydrophobic wedge is depicted in green, the helical arch in red, the H2TH domain in light blue and the rest of the crystallized structure in dark blue. (b) Structure of Rad2 with dsDNA. Same color code as in (a). Calcium is depicted in light green and potassium in purple. The helix 12b is named as referred by Tsutakawa *et al.* (2011). The spacer region omitted in the crystal structure is depicted as a yellow dashed line.

three structural domains to recognize the double-stranded portion of the DNA but it does not specifically recognize the single-stranded portion of the duplex. The structures reveal that the protein recognizes a single 5'-nucleotide of the single-stranded portion of the DNA and a 3'-phosphate group of the ds/ssDNA junction. The helix two turn helix (H2TH) motif including a Rad2-specific α -helix 12b and a hydrophobic wedge, which binds the first base pair of the double-stranded portion of the DNA, is responsible for binding to DNA. The helical arch may assume a different structure in Rad2 because its active site is more accessible compared with related enzymes such as FEN1 and EXO1 (Mietus *et al.* 2014). This feature allows the protein to form a gap and create an exit route from the active site. As expected, the catalytic core of Rad2 is quite comparable with other flap nucleases (Orans *et al.* 2011). The protein has an oblong shape with its active site situated in the middle. A twisted central β -sheet composed of seven strands flanked by α -helices build the central core of the protein (Fig. 16b). Gln37 forms a hydrogen bond with the base of the cleaved strand and represents an important residue of substrate recognition by Rad2. The H2TH module is responsible for the key protein–DNA interactions through its Rad2-specific helix α 12b and the hydrophobic wedge contacting the ss/dsDNA junction.



9. XPV

The repair of lesions by NER is a powerful method to clean the genome from the plethora of lesions that are generated daily. NER is hence the prime pathway that allows cells to cope with DNA damage. Because lesions can occur at any time in cellular DNA, even when the NER complex cannot assemble properly, e.g. while the cell is undergoing cell division, a second system has evolved that allows cells to cope with lesions. This system does not repair lesions but is able to give cells the ability to tolerate lesions, so that repair can be shifted to later times. This lesion tolerance process (called TLS) requires the action of special lesion tolerance polymerases. Among them the polymerase η (XPV) is of prime importance, because it provides tolerance of dinucleotide lesions such as CPDs and cisplatin adducts, which are the prime repair targets for NER. XPV (Polymerase η) is a Y-family DNA polymerase responsible for TLS enabling replication through bulky DNA adducts like cisplatin, UV-induced CPDs and 6–4PP as well as acetylaminofluorene adducts and oxidative lesions (like 8-oxo-G) (Alt *et al.* 2007; Biertumpfel *et al.* 2010; Haracska *et al.* 2000; Reissner *et al.* 2010; Schorr *et al.* 2010b; Silverstein *et al.* 2010a, b). The human protein consists of 713 amino acids with a size of 78 kDa (Masutani *et al.* 1999a). Indeed, mutation in the XPV gene leads to just another form of XP called the variant type also associated with a high risk of developing skin cancer (Johnson *et al.* 2000; Lehmann *et al.* 1975; Masutani *et al.* 1999b). Lesion tolerance, however, has a major drawback. Because the polymerases are able to read through lesions they also have a strongly reduced replication fidelity, which gives rise to mutations. That is why repair is always the first and preferred defense pathway.

9.1 Crystal structure of the yeast XPV (Rad30) protein

Kondratyuk *et al.* (2001) solved the crystal structure of the catalytic core (1–513 amino acids) of the *S. cerevisiae* apo-enzyme (Fig. 17b). The catalytic core possesses the same bypass activity as the full-length protein (1–632 amino acids). The structure shows that the polymerase possesses a right-hand shape built up by four domains: thumb, fingers, PAD-domain (polymerase-associated domain also named ‘little fingers’ (LF)) and the palm domain (Fig. 17a). Characterized by a wide open and possibly flexible active site, the polymerase is able to replicate over bulky dinucleotide lesions, however, at the cost of misincorporation and only a low processivity in contrast to high-fidelity polymerases. Because of the helix deforming action of lesions like intrastrand crosslinks (CPD and 1,2-GG) their bypass by high-fidelity DNA polymerases is impossible. Being able to bypass cisplatin crosslinks, XPV enables cells to gain resistance against cisplatin-based chemotherapy.

9.2 Crystal structure of the yeast XPV (Rad30) in complex with 1,2-GG cisplatin

Alt *et al.* (2007) solved the crystal structure of XPV replicating over 1,2-GG containing DNA. This structure provided insights into the molecular mechanism that allows replication of XPV across strongly distorting DNA lesions. 1,2-GG and 1,3-GTG cisplatin adducts are formed in a typical cancer therapy with cisplatin. In the structure, XPV accommodates the full 1,2-GG adduct in its active site where it properly base pairs the 3' G of the 1,2-GG lesion with a dCTP (Fig. 18). For the nucleotidyl transfer, the DNA rotates into an active conformation, supported by hydrogen bonding of the template base to the dCTP. The bypass of the 3'-dG of 1,2-GG is more efficient and accurate than the bypass of the 5'-dG. In the palm domain, Asp30 and Asp155 coordinate the divalent cations in the active site and Glu156 is responsible for the catalysis. Glu39 plays a structural role too, because it maintains the integrity of the fingers domain. Arg73 coordinates the incoming dNTP for the lesion bypass step and stacks on top of the base moiety. Two magnesium ions are coordinated by the catalytically essential residues.

9.3 Crystal structure of the yeast XPV (Rad30) in complex with the CPD lesion

Silverstein *et al.* (2010b) determined the crystal structure of yeast Rad30 in complex with a *cis-syn* T–T dimer lesion. These authors argue that crystal contacts between the palm and the PAD domain of two neighboring proteins in the crystal lattice influence the binding of the protein to the DNA. The group performed a double-mutation (K140A, S144W) to eliminate these contacts, and indeed crystals were obtained with a different space group. Deeper analysis of the structure shows that the finger domain is a little more closed but the situation in the active site is almost unchanged compared with the yeast structure (Alt *et al.* 2007). Figure 19 shows the XPV structure with its different domains, the containing T–T dimer and the incoming dATP. Again the large active site accommodates the dinucleotide lesion very similar to what was already shown by Alt *et al.* for the cisplatin lesion. The 3'T of the dimer interacts with the incoming dATP by Watson–Crick hydrogen bonding. The 5'T of the lesion builds van der Waals contacts with Met74 and Arg73 of the finger domain. The active site residues of the palm domain (Asp30, Asp155 and Glu156) catalyze the nucleotidyl transfer reaction. By mutational analysis and biochemical assays Silverstein *et al.* could show that the residues Gln55, Arg73 and Met74 interact with the T–T dimer to keep the lesion in the active site.

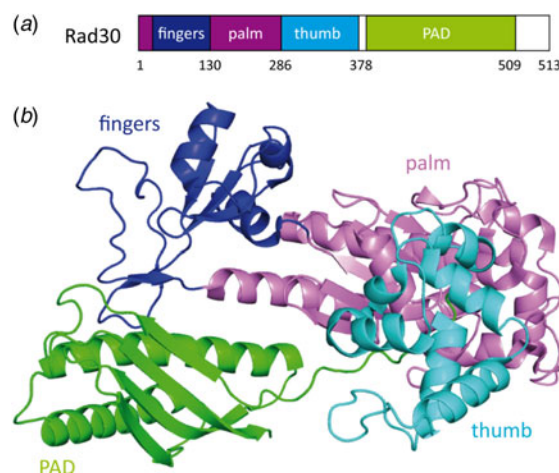


Fig. 17. (a) Structural domains of yeast XPV (Rad30) (Silverstein *et al.* 2010b). (b) Structure of the catalytic domain of *S. cerevisiae* XPV showing the four domains (PDB code: 1JIH) (Trincao *et al.* 2001). The fingers domain is depicted in dark blue, the PAD domain in green, the ThM in cyan and the palm domain in purple.

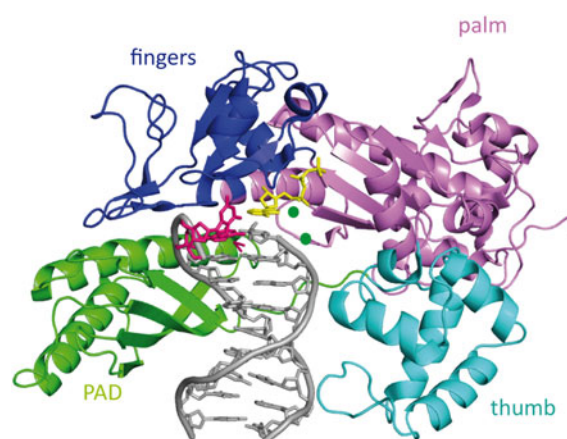


Fig. 18. Crystal structure of Rad30 in complex with 1,2-GG cisplatin adduct (PDB code: 2R8 K) (Alt *et al.* 2007). Rad30 has right-hand structure containing four domains: palm (violet), thumb (cyan), finger domain (dark blue) and the additional PAD domain (green) which is unique for Y-family polymerases. The cisplatin adduct is depicted in pink, the dCTP in yellow and the two metal ions in green.

9.4 Crystal structure of the human XPV in complex with the CPD lesion

Biertumpfel *et al.* (2010) reported the crystal structure of the catalytic core (residues 1–432) of human polymerase η at four consecutive steps during DNA synthesis through *cis-syn* CPD lesion (Fig. 20b). They could show on a molecular level how human XPV efficiently bypasses the lesion in TLS. Human XPV, similar to its yeast homolog, consists of four domains: palm, PAD (little fingers), thumb and the fingers (Fig. 20a). The active site is situated in the palm domain and the DNA is bound between the thumb and the finger domains. Acting as a molecular splint the polymerase keeps the damaged DNA straight and rigid in its B-form conformation in contrast to the bent and unwound structures of CPD containing DNA with other repair enzymes or alone. The β -strand in the little finger is nearly parallel to the template strand and hydrogen bonding of the main chain amides to the phosphate supports this conformation. Again, the most prominent feature is the large active site of the protein already reported by Alt *et al.* (2007), which is able to accommodate two bases of the template strand instead of one. The 3' thymine of CPD is base paired with the incoming dATP, while the 5' thymine is turned and moved closer to the finger domain which leads to tighter interactions than the ones observed with undamaged DNA. Furthermore, the two residues Arg61 and Glu38 that are conserved among the XPV homologs create specific hydrogen bonds with the incoming dATP and the CPD lesion. Again, the human XPV shows a more closed finger domain, which contacts the replicating base pair more intensively. Based on all these structures we can therefore assume that the finger domain is quite flexible and that the more closed structure is the conformation the protein assumes in the processive mode.

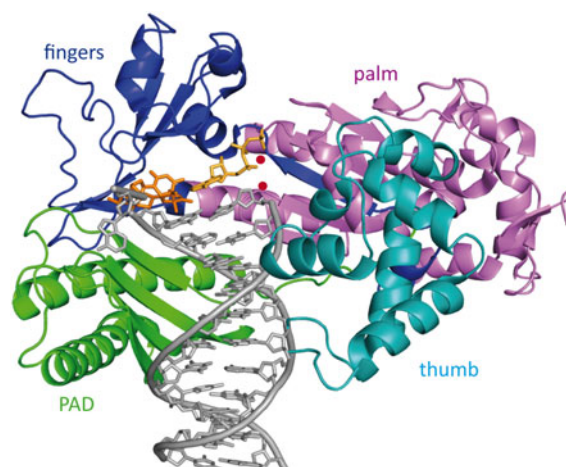


Fig. 19. Structure of Rad30 with a CPD lesion. The *cis-syn* T-T dimer is depicted in orange and the incoming dATP is shown in yellow (PDB code: 3MFI) (Silverstein *et al.* 2010b). The four domains are depicted with different colors: fingers (blue), palm (violet), PAD (green) and thumb (light blue). The two Mg^{2+} ions in the active site are colored in red.

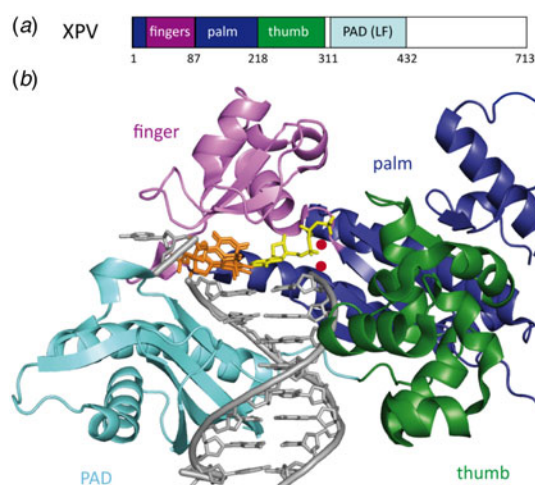


Fig. 20. (a) Schematic representation of human XPV showing the different domains. The finger domain is depicted in purple, the palm domain in dark blue, the ThM in green and PAD domain (little fingers) in cyan. (b) Crystal structure of human XPV in complex with CPD damaged DNA (PDB code: 3MR3) (Biertumpfel *et al.* 2010). Same color code as in (a). The CPD lesion is shown in orange and the dAMPNPP (2'-deoxy-adenosine-5'-[(α,β)-imido]-triphosphate) is shown in yellow. The two Mg^{2+} ions of the active site are depicted in red.

Further crystal structures of human and yeast XPV could be resolved with different DNA damages such as 1,3-GTG cisplatin or 8-oxoG (Reissner *et al.* 2010; Silverstein *et al.* 2010a). Schorr & Carell (2010a) crystallized the *S. cerevisiae* XPV (Rad30) with an AAF-dG lesion and they proposed a model for lesion bypass. *In vivo* experiments show that Pol η indeed replicates through acetylated aromatic amine lesions such as AAF-dG (Bresson & Fuchs, 2002). *In vitro* studies revealed that both human and yeast XPV base pair the AAF-dG lesion correctly with dC (Bresson & Fuchs, 2002; Livneh, 2001; Masutani *et al.* 2000; Matsuda *et al.* 2000) and are able to catalyze the extension from the lesion, however, with reduced efficiency (Masutani *et al.* 2000). In addition, XPV also incorporates a correct dCTP opposite the N^2 -AAF-dG adduct (Yasui *et al.* 2004). It has also been shown that XPV is responsible for the occurrence of frameshift mutations in response to the presence of aromatic amine lesions during replication (Bresson & Fuchs, 2002; Cordonnier *et al.* 1999).

10. Conclusion

In this review, we presented the most recent eukaryotic structures of the XP proteins in complex with their DNA substrate. When no co-crystal structures with the DNA substrate were obtained we discussed the models of protein–DNA interaction



that were proposed. All these structures led to a dramatic increase of our understanding of the role of the individual XP proteins in DNA damage repair and tolerance. The structures allow to determine binding partner regions and to elucidate on a molecular level the interaction between the proteins and their DNA substrates. Now, further studies have to be done in order to better understand the correlation between genetic defects and the role of the various XP proteins in NER and TLS. We also need further human structural information and most importantly structures of the XP proteins in functional complexes with the other involved NER proteins. This would allow deciphering of the protein–protein interactions on the DNA damage site and it may shed light on the correlation to genetic disorders. Understanding the mechanism of NER and TLS in correlation with human genetic disorders unveils the biological basis for oncogenesis and aging. Last but not least, proteomics data are needed to uncover the functional relations of the proteins involved in replication, transcription and repair.

Acknowledgements

We thank the Deutsche Forschungsgemeinschaft SFB-646, SFB-749 and the excellence cluster CIPS^M for financial support. We also thank Dr. Jochen Kuper, Wolfgang Koelmel, Dr. Dorothea Matschkal, Dr. Iacovos N. Michaelides and Dr. Sabine Schneider for critical reading of the manuscript. We thank Dr. Markus Mueller for helpful advice.

References

- ABDULRAHMAN, W., ILTIS, I., RADU, L., BRAUN, C., MAGLOTT-ROTH, A., GIRAUDON, C., EGLY, J. M. & POTERSZMAN, A. (2013). ARCH domain of XPD, an anchoring platform for CAK that conditions TFIIH DNA repair and transcription activities. *Proceedings of the National Academy of Sciences of the United States of America* **110**, E633–E642.
- ABOUSSEKHRA, A., BIGGERSTAFF, M., SHIVJI, M. K., VILPO, J. A., MONCOLLIN, V., PODUST, V. N., PROTIC, M., HUBSCHER, U., EGLY, J. M. & WOOD, R. D. (1995). Mammalian DNA nucleotide excision repair reconstituted with purified protein components. *Cell* **80**, 859–868.
- ALT, A., LAMMENS, K., CHIOCCINI, C., LAMMENS, A., PIECK, J. C., KUCH, D., HOPFNER, K. P. & CARELL, T. (2007). Bypass of DNA lesions generated during anticancer treatment with cisplatin by DNA polymerase ϵ . *Science* **318**, 967–970.
- ARAKI, M., MASUTANI, C., TAKEMURA, M., UCHIDA, A., SUGASAWA, K., KONDOH, J., OHKUMA, Y. & HANAOKA, F. (2001). Centrosome protein centrin 2/caltractin 1 is part of the xeroderma pigmentosum group C complex that initiates global genome nucleotide excision repair. *The Journal of Biological Chemistry* **276**, 18665–18672.
- ARAUJO, S. J., NIGG, E. A. & WOOD, R. D. (2001). Strong functional interactions of TFIIH with XPC and XPG in human DNA nucleotide excision repair, without a preassembled repairosome. *Molecular and Cellular Biology* **21**, 2281–2291.
- ARAUJO, S. J., TIRODE, F., COIN, F., POSPIECH, H., SYVAOJA, J. E., STUCKI, M., HUBSCHER, U., EGLY, J. M. & WOOD, R. D. (2000). Nucleotide excision repair of DNA with recombinant human proteins: definition of the minimal set of factors, active forms of TFIIH, and modulation by CAK. *Genes & Development* **14**, 349–359.
- ASAHINA, H., KURAOKA, I., SHIRAKAWA, M., MORITA, E. H., MIURA, N., MIYAMOTO, I., OHTSUKA, E., OKADA, Y. & TANAKA, K. (1994). The XPA protein is a zinc metalloprotein with an ability to recognize various kinds of DNA damage. *Nature Structural & Molecular Biology* **315**, 229–237.
- BATTY, D., RAPIC'-OTRIN, V., LEVINE, A. S. & WOOD, R. D. (2000a). Stable binding of human XPC complex to irradiated DNA confers strong discrimination for damaged sites. *Journal of Molecular Biology* **300**, 275–290.
- BATTY, D. P. & WOOD, R. D. (2000b). Damage recognition in nucleotide excision repair of DNA. *Gene* **241**, 193–204.
- BERNARDES DE JESUS, B. M., BJORAS, M., COIN, F. & EGLY, J. M. (2008). Dissection of the molecular defects caused by pathogenic mutations in the DNA repair factor XPC. *Molecular and Cellular Biology* **28**, 7225–7235.
- BERNEBURG, M. & LEHMANN, A. R. (2001). Xeroderma pigmentosum and related disorders: defects in DNA repair and transcription. *Advances in Genetics* **43**, 71–102.
- BIERTUMPFEL, C., ZHAO, Y., KONDO, Y., RAMON-MAIQUES, S., GREGORY, M., LEE, J. Y., MASUTANI, C., LEHMANN, A. R., HANAOKA, F. & YANG, W. (2010). Structure and mechanism of human DNA polymerase ϵ . *Nature* **465**, 1044–1048.
- BOOTSMA, D. & HOEIJMAKERS, J. H. (1993). DNA repair. Engagement with transcription. *Nature* **363**, 114–115.
- BRESSON, A. & FUCHS, R. P. (2002). Lesion bypass in yeast cells: Pol ϵ participates in a multi-DNA polymerase process. *The EMBO Journal* **21**, 3881–3887.
- BROOKMAN, K. W., LAMERDIN, J. E., THELEN, M. P., HWANG, M., REARDON, J. T., SANCAR, A., ZHOU, Z. Q., WALTER, C. A., PARRIS, C. N. & THOMPSON, L. H. (1996). ERCC4 (XPF) encodes a human nucleotide excision repair protein with eukaryotic recombination homologs. *Molecular and Cellular Biology* **16**, 6553–6562.
- BROOKS, P. J., WISE, D. S., BERRY, D. A., KOSMOSKI, J. V., SMERDON, M. J., SOMERS, R. L., MACKIE, H., SPOONDE, A. Y., ACKERMAN, E. J., COLEMAN, K., TARONE, R. E. & ROBBINS, J. H. (2000). The oxidative DNA lesion 8,5'-(S)-cyclo-2'-deoxyadenosine is repaired by the nucleotide excision repair pathway and blocks gene expression in mammalian cells. *The Journal of Biological Chemistry* **275**, 22355–22362.
- BUCHKO, G. W., NI, S., THRALL, B. D. & KENNEDY, M. A. (1998). Structural features of the minimal DNA binding domain (M98-F219) of human nucleotide excision repair protein XPA. *Nucleic Acids Research* **26**, 2779–2788.

- BUCHKO, G. W., TUNG, C. S., MCAETER, K., ISERN, N. G., SPICER, L. D. & KENNEDY, M. A. (2001). DNA-XPA interactions: a ^{31}P NMR and molecular modeling study of dCCAATAACC association with the minimal DNA-binding domain (M98-F219) of the nucleotide excision repair protein XPA. *Nucleic Acids Research* **29**, 2635–2643.
- BUECHNER, C. N., HEIL, K., MICHELS, G., CARELL, T., KISKER, C. & TESSMER, I. (2014). Strand-specific recognition of DNA damages by XPD provides insights into nucleotide excision repair substrate versatility. *The Journal of Biological Chemistry* **289**, 3613–3624.
- BURNS, J. L., GUZDER, S. N., SUNG, P., PRAKASH, S. & PRAKASH, L. (1996). An affinity of human replication protein A for ultraviolet-damaged DNA. *The Journal of Biological Chemistry* **271**, 11607–11610.
- BUSCHTA-HEDAYAT, N., BUTERIN, T., HESS, M. T., MISSURA, M. & NAEGLI, H. (1999). Recognition of nonhybridizing base pairs during nucleotide excision repair of DNA. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 6090–6095.
- BUTERIN, T., HESS, M. T., LUNEVA, N., GEACINTOV, N. E., AMIN, S., KROTH, H., SEIDEL, A. & NAEGLI, H. (2000). Unrepaired Fjord region polycyclic aromatic hydrocarbon-DNA adducts in ras codon 61 mutational hot spots. *Cancer Research* **60**, 1849–1856.
- CAMENISCH, U., DIP, R., SCHUMACHER, S. B., SCHULER, B. & NAEGLI, H. (2006). Recognition of helical kinks by xeroderma pigmentosum group A protein triggers DNA excision repair. *Nature Structural and Molecular Biology* **13**, 278–284.
- CAMENISCH, U. & NAEGLI, H. (2008). XPA gene, its product and biological roles. *Advances in Experimental Medicine and Biology* **637**, 28–38.
- CAMENISCH, U., TRAUTLEIN, D., CLEMENT, F., FEI, J., LEITENSTORFER, A., FERRANDO-MAY, E. & NAEGLI, H. (2009). Two-stage dynamic DNA quality check by xeroderma pigmentosum group C protein. *EMBO Journal* **28**, 2387–2399.
- CHEN, X., VELMURUGU, Y., ZHENG, G., PARK, B., SHIM, Y., KIM, Y., LIU, L., VAN HOUTEN, B., HE, C., ANSARIA, A. & MIN, J. (2015). Kinetic gating mechanism of DNA damage recognition by Rad4/XPC. *Nature Communications* **6**, 5849.
- CHU, G. & CHANG, E. (1988). Xeroderma pigmentosum group E cells lack a nuclear factor that binds to damaged DNA. *Science* **242**, 564–567.
- CITTERIO, E., VAN DEN BOOM, V., SCHNITZLER, G., KANAAR, R., BONTE, E., KINGSTON, R. E., HOEIJMAKERS, J. H. & VERMEULEN, W. (2000). ATP-dependent chromatin remodeling by the Cockayne syndrome B DNA repair-transcription-coupling factor. *Molecular and Cellular Biology* **20**, 7643–7653.
- CLEAVER, J. E. (2000). Common pathways for ultraviolet skin carcinogenesis in the repair and replication defective groups of xeroderma pigmentosum. *Journal of Dermatological Science* **23**, 1–11.
- CLEAVER, J. E., LAM, E. T. & REVET, I. (2009). Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity. *Nature Reviews Genetics* **10**, 756–768.
- CLEAVER, J. E. & STATES, J. C. (1997). The DNA damage-recognition problem in human and other eukaryotic cells: the XPA damage binding protein. *Biochemical Journal* **328**, 1–12.
- COIN, F., MARINONI, J. C., RODOLFO, C., FRIBOURG, S., PEDRINI, A. M. & EGLY, J. M. (1998). Mutations in the XPD helicase gene result in XP and TTD phenotypes, preventing interaction between XPD and the p44 subunit of TFIIH. *Nature Genetics* **20**, 184–188.
- COIN, F., OKSENYCH, V. & EGLY, J. M. (2007). Distinct roles for the XPB/p52 and XPD/p44 subcomplexes of TFIIH in damaged DNA opening during nucleotide excision repair. *Molecular Cell* **26**, 245–256.
- COIN, F., OKSENYCH, V., MOCQUET, V., GROH, S., BLATTNER, C. & EGLY, J. M. (2008). Nucleotide excision repair driven by the dissociation of CAK from TFIIH. *Molecular Cell* **31**, 9–20.
- CONSTANTINO, A., GUNZ, D., EVANS, E., LALLE, P., BATES, P. A., WOOD, R. D. & CLARKSON, S. G. (1999). Conserved residues of human XPG protein important for nuclease activity and function in nucleotide excision repair. *The Journal of Biological Chemistry* **274**, 5637–5648.
- CORDONNIER, A. M., LEHMANN, A. R. & FUCHS, R. P. (1999). Impaired translesion synthesis in xeroderma pigmentosum variant extracts. *Molecular and Cellular Biology* **19**, 2206–2211.
- DAS, D., FOLKERS, G. E., VAN DIJK, M., JASPERS, N. G., HOEIJMAKERS, J. H., KAPTEIN, R. & BOELEN, R. (2012). The structure of the XPF-ssDNA complex underscores the distinct roles of the XPF and ERCC1 helix-hairpin-helix domains in ss/ds DNA recognition. *Structure* **20**, 667–675.
- DAS, D., TRIPSANES, K., JASPERS, N. G., HOEIJMAKERS, J. H., KAPTEIN, R., BOELEN, R. & FOLKERS, G. E. (2008). The HhH domain of the human DNA repair protein XPF forms stable homodimers. *Proteins* **70**, 1551–1563.
- DE LAAT, W. L., APPELDOORN, E., SUGASAWA, K., WETERINGS, E., JASPERS, N. G. J. & HOEIJMAKERS, J. H. J. (1998). DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. *Genes & Development* **12**, 2598–2609.
- DE LAAT, W. L., JASPERS, N. G. & HOEIJMAKERS, J. H. (1999). Molecular mechanism of nucleotide excision repair. *Genes & Development* **13**, 768–785.
- DE WEERD-KASTELEIN, E. A., KEIJZER, W. & BOOTSMA, D. (1972). Genetic heterogeneity of xeroderma pigmentosum demonstrated by somatic cell hybridization. *Nature: New Biology* **238**, 80–83.
- DEMPLE, B. & HARRISON, L. (1994). Repair of oxidative damage to DNA: enzymology and biology. *Annual Review of Biochemistry* **63**, 915–948.
- DONAHUE, B. A., YIN, S., TAYLOR, J. S., REINES, D. & HANAWALT, P. C. (1994). Transcript cleavage by RNA polymerase II arrested by a cyclobutane pyrimidine dimer in the DNA template. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 8502–8506.
- DUBAELE, S., PROIETTI DE SANTIS, L., BIENSTOCK, R. J., KERIEL, A., STEFANINI, M., VAN HOUTEN, B. & EGLY, J. M. (2003). Basal transcription defect discriminates between xeroderma pigmentosum and trichothiodystrophy in XPD patients. *Molecular Cell* **11**, 1635–1646.
- DUNAND-SAUTHER, I., HOHL, M., THOREL, F., JAQUIER-GUBLER, P., CLARKSON, S. G. & SCHARER, O. D. (2005). The spacer region of XPG mediates recruitment to nucleotide excision repair complexes and determines substrate specificity. *The Journal of Biological Chemistry* **280**, 7030–7037.
- EDENBERG, H. & HANAWALT, P. (1972). Size of repair patches in the DNA of ultraviolet-irradiated HeLa cells. *Biochimica et Biophysica Acta* **272**, 361–372.

- EL-MAHDY, M. A., ZHU, Q., WANG, Q. E., WANI, G., PRAETORIUS-IBBA, M., WANI, A. A. (2006). Cullin 4A-mediated proteolysis of DDB2 protein at DNA damage sites regulates *in vivo* lesion recognition by XPC. *Journal of Biological Chemistry* **281**, 13404–13411.
- EVANS, E., FELLOWS, J., COFFER, A. & WOOD, R. D. (1997a). Open complex formation around a lesion during nucleotide excision repair provides a structure for cleavage by human XPG protein. *The EMBO Journal* **16**, 625–638.
- EVANS, E., MOGGS, J. G., HWANG, J. R., EGLY, J. M. & WOOD, R. D. (1997b). Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. *The EMBO Journal* **16**, 6559–6573.
- FAGBEMI, A. F., ORELLI, B. & SCHARER, O. D. (2011). Regulation of endonuclease activity in human nucleotide excision repair. *DNA Repair* **10**, 722–729.
- FAN, L., ARVAL, A. S., COOPER, P. K., IWAI, S., HANAOKA, F. & TAINER, J. A. (2006). Conserved XPB core structure and motifs for DNA unwinding: implications for pathway selection of transcription or excision repair. *Molecular Cell* **22**, 27–37.
- FAN, L., FUSS, J. O., CHENG, Q. J., ARVAL, A. S., HAMMEL, M., ROBERTS, V. A., COOPER, P. K. & TAINER, J. A. (2008). XPD helicase structures and activities: insights into the cancer and aging phenotypes from XPD mutations. *Cell* **133**, 789–800.
- FELTES, B. C. & BONATTO, D. (2015). Overview of xeroderma pigmentosum proteins architecture, mutations and post-translational modifications. *Mutation Research/Reviews in Mutation Research* **763**, 306–320.
- FISCHER, E. S., SCRIMA, A., BOHM, K., MATSUMOTO, S., LINGARAJU, G. M., FATY, M., YASUDA, T., CAVADINI, S., WAKASUGI, M., HANAOKA, F., IWAI, S., GUT, H., SUGASAWA, K. & THOMA, N. H. (2011). The molecular basis of CRL4DDB2/CSA ubiquitin ligase architecture, targeting, and activation. *Cell* **147**, 1024–1039.
- FITCH, M. E., NAKAJIMA, S., YASUI, A. & FORD, J. M. (2003). *In vivo* recruitment of XPC to UV-induced cyclobutane pyrimidine dimers by the DDB2 gene product. *The Journal of Biological Chemistry* **278**, 46906–46910.
- FRIEDBERG, E. C. (1995). *DNA Repair and Mutagenesis*. Washington, D.C: ASM Press.
- FRIEDBERG, E. C. (2001). How nucleotide excision repair protects against cancer. *Nature Review Cancer* **1**, 22–33.
- FRIEDBERG, E. C. (2005). Suffering in silence: the tolerance of DNA damage. *Nature Reviews Molecular Cell Biology* **6**, 943–953.
- FUJIWARA, Y., MASUTANI, C., MIZUKOSHI, T., KONDO, J., HANAOKA, F. & IWAI, S. (1999). Characterization of DNA recognition by the human UV-damaged DNA-binding protein. *The Journal of Biological Chemistry* **274**, 20027–20033.
- FUSS, J. O. & TAINER, J. A. (2011). XPB and XPD helicases in TFIIH orchestrate DNA duplex opening and damage verification to coordinate repair with transcription and cell cycle via CAK kinase. *DNA Repair* **10**, 697–713.
- GEACINTOV, N. E., BRODY, S., BUTERIN, T., NAEGELI, H., WU, M., YAN, S. & PATEL, D. J. (2002). Thermodynamic and structural factors in the removal of bulky DNA adducts by the nucleotide excision repair machinery. *Biopolymers* **65**, 202–210.
- GILLET, L. C., ALZEER, J. & SCHARER, O. D. (2005). Site-specific incorporation of N-(deoxyguanosin-8-yl)-2-acetylaminofluorene (dG-AAF) into oligonucleotides using modified ‘ultra-mild’ DNA synthesis. *Nucleic Acids Research* **33**, 1961–1969.
- GILLET, L. C. & SCHARER, O. D. (2006). Molecular mechanisms of mammalian global genome nucleotide excision repair. *Chemical Reviews* **106**, 253–276.
- GROISMAN, R., POLANOWSKA, J., KURAOKA, I., SAWADA, J., SAJO, M., DRAPKIN, R., KISSELEV, A. F., TANAKA, K., NAKATANI, Y. (2003). The ubiquitin ligase activity in the DDB2 and CSA complexes is differentially regulated by the COP9 signalosome in response to DNA damage. *Cell* **113**, 357–367.
- GUNZ, D., HESS, M. T. & NAEGELI, H. (1996). Recognition of DNA adducts by human nucleotide excision repair. Evidence for a thermodynamic probing mechanism. *The Journal of Biological Chemistry* **271**, 25089–25098.
- GUZDER, S. N., SUNG, P., BAILLY, V., PRAKASH, L. & PRAKASH, S. (1994). RAD25 is a DNA helicase required for DNA repair and RNA polymerase II transcription. *Nature* **369**, 578–581.
- HANAWALT, P. C. & SPIVAK, G. (2008). Transcription-coupled DNA repair: two decades of progress and surprises. *Nature Reviews Molecular Cell Biology* **9**, 958–970.
- HARACSKA, L., YU, S. L., JOHNSON, R. E., PRAKASH, L. & PRAKASH, S. (2000). Efficient and accurate replication in the presence of 7,8-dihydro-8-oxoguanine by DNA polymerase ϵ . *Nature Genetics* **25**, 458–461.
- HE, Z., HENRICKSEN, L. A., WOLD, M. S. & INGLES, C. J. (1995). RPA involvement in the damage-recognition and incision steps of nucleotide excision repair. *Nature* **374**, 566–569.
- HEBRA, F. & KAPOSI, M. (1874). On diseases of the skin including the exanthemata. *The New Sydenham Society* **61**, 252–258.
- HESS, M. T., GUNZ, D., LUNEVA, N., GEACINTOV, N. E. & NAEGELI, H. (1997a). Base pair conformation-dependent excision of benzo[a]pyrene diol epoxide-guanine adducts by human nucleotide excision repair enzymes. *Molecular and Cellular Biology* **17**, 7069–7076.
- HESS, M. T., GUNZ, D. & NAEGELI, H. (1996a). A repair competition assay to assess recognition by human nucleotide excision repair. *Nucleic Acids Research* **24**, 824–828.
- HESS, M. T., SCHWITTER, U., PETRETTA, M., GIESE, B. & NAEGELI, H. (1996b). Site-specific DNA substrates for human excision repair: comparison between deoxyribose and base adducts. *Chemical Biology* **3**, 121–128.
- HESS, M. T., SCHWITTER, U., PETRETTA, M., GIESE, B. & NAEGELI, H. (1997b). Bipartite substrate discrimination by human nucleotide excision repair. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 6664–6669.
- HESS, N. J., BUCHKO, G. W., CONRADSON, S. D., ESPINOSA, F. J., NI, S., THRALL, B. D. & KENNEDY, M. A. (1998). Human nucleotide excision repair protein XPA: extended X-ray absorption fine-structure evidence for a metal-binding domain. *Protein Science* **7**, 1970–1975.
- HEY, T., LIPPS, G., SUGASAWA, K., IWAI, S., HANAOKA, F. & KRAUSS, G. (2002). The XPC-HR23B complex displays high affinity and specificity for damaged DNA in a true-equilibrium fluorescence assay. *Biochemistry* **41**, 6583–6587.
- HILARIO, E., LI, Y., NOBUMORI, Y., LIU, X. & FAN, L. (2013). Structure of the C-terminal half of human XPB helicase and the impact of the disease-causing mutation XP11BE. *Acta Crystallographica D: Biological Crystallography* **69**, 237–246.



- HOEIJMAKERS, J. H. (2001). Genome maintenance mechanisms for preventing cancer. *Nature* **411**, 366–374.
- HOHL, M., DUNAND-SAUTHER, I., STARESINCIC, L., JAQUIER-GUBLER, P., THOREL, F., MODESTI, M., CLARKSON, S. G. & SCHARER, O. D. (2007). Domain swapping between FEN-1 and XPG defines regions in XPG that mediate nucleotide excision repair activity and substrate specificity. *Nucleic Acids Research* **35**, 3053–3063.
- HOHL, M., THOREL, F., CLARKSON, S. G. & SCHARER, O. D. (2003). Structural determinants for substrate binding and catalysis by the structure-specific endonuclease XPG. *The Journal of Biological Chemistry* **278**, 19500–19508.
- HOOGSTRATEN, D., NIGG, A. L., HEATH, H., MULLENDERS, L. H., VAN DRIEL, R., HOEIJMAKERS, J. H., VERMEULEN, W. & HOUTSMULLER, A. B. (2002). Rapid switching of TFIIH between RNA polymerase I and II transcription and DNA repair *in vivo*. *Molecular Cell* **10**, 1163–1174.
- HOSFIELD, D. J., MOL, C. D., SHEN, B. & TAINER, J. A. (1998). Structure of the DNA repair and replication endonuclease and exonuclease FEN-1: coupling DNA and PCNA binding to FEN-1 activity. *Cell* **95**, 135–146.
- HOULE, J. F. & FRIEDBERG, E. C. (1999). The *Drosophila* ortholog of the human XPG gene. *Gene* **234**, 353–360.
- HOUTSMULLER, A. B., RADEMAKERS, S., NIGG, A. L., HOOGSTRATEN, D., HOEIJMAKERS, J. H. & VERMEULEN, W. (1999). Action of DNA repair endonuclease ERCC1/XPF in living cells. *Science* **284**, 958–961.
- HUANG, J. C., HSU, D. S., KAZANTSEV, A. & SANCAR, A. (1994). Substrate spectrum of human excinuclease: repair of abasic sites, methylated bases, mismatches, and bulky adducts. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 12213–12217.
- HWANG, B. J., FORD, J. M., HANAWALT, P. C. & CHU, G. (1999). Expression of the p48 xeroderma pigmentosum gene is p53-dependent and is involved in global genomic repair. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 424–428.
- HWANG, B. J., TOERING, S., FRANCKE, U. & CHU, G. (1998a). p48 Activates a UV-damaged-DNA binding factor and is defective in xeroderma pigmentosum group E cells that lack binding activity. *Molecular and Cellular Biology* **18**, 4391–4399.
- HWANG, K. Y., BAEK, K., KIM, H. Y. & CHO, Y. (1998b). The crystal structure of flap endonuclease-1 from *Methanococcus jannaschii*. *Nature Structural & Molecular Biology* **5**, 707–713.
- KEGAMI, T., KURAOKA, I., SAJO, M., KODO, N., KYOGOKU, Y., MORIKAWA, K., TANAKA, K. & SHIRAKAWA, M. (1998). Solution structure of the DNA- and RPA-binding domain of the human repair factor XPA. *Nature Structural & Molecular Biology* **5**, 701–706.
- ISAACS, R. J. & SPIELMANN, H. P. (2004). A model for initial DNA lesion recognition by NER and MMR based on local conformational flexibility. *DNA Repair* **3**, 455–464.
- IYER, N., REAGAN, M. S., WU, K. J., CANAGARAJAH, B. & FRIEDBERG, E. C. (1996). Interactions involving the human RNA polymerase II transcription/nucleotide excision repair complex TFIIH, the nucleotide excision repair protein XPG, and Cockayne syndrome group B (CSB) protein. *Biochemistry* **35**, 2157–2167.
- JOHNSON, R. E., WASHINGTON, M. T., HARACKA, L., PRAKASH, S. & PRAKASH, L. (2000). Eukaryotic polymerases ι and ζ act sequentially to bypass DNA lesions. *Nature* **406**, 1015–1019.
- JONES, C. J. & WOOD, R. D. (1993). Preferential binding of the xeroderma pigmentosum group A complementing protein to damaged DNA. *Biochemistry* **32**, 12096–12104.
- KAMIUCHI, S., SAJO, M., CITTERIO, E., DE JAGER, M., HOEIJMAKERS, J. H. & TANAKA, K. (2002). Translocation of Cockayne syndrome group A protein to the nuclear matrix: possible relevance to transcription-coupled DNA repair. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 201–206.
- KATSUMI, S., KOBAYASHI, N., IMOTO, K., NAKAGAWA, A., YAMASHINA, Y., MURAMATSU, T., SHIRAI, T., MIYAGAWA, S., SUGIURA, S., HANAOKA, F., MATSUNAGA, T., NIKAI, O. & MORI, T. (2001). *In situ* visualization of ultraviolet-light-induced DNA damage repair in locally irradiated human fibroblasts. *Journal of Investigative Dermatology* **117**, 1156–1161.
- KEENEY, S., CHANG, G. J. & LINN, S. (1993). Characterization of a human DNA damage binding protein implicated in xeroderma pigmentosum E. *The Journal of Biological Chemistry* **268**, 21293–21300.
- KIM, T. K., EBRIGHT, R. H. & REINBERG, D. (2000). Mechanism of ATP-dependent promoter melting by transcription factor IIIH. *Science* **288**, 1418–1422.
- KOBAYASHI, T., TAKEUCHI, S., SAJO, M., NAKATSU, Y., MORIOKA, H., OTSUKA, E., WAKASUGI, M., NIKAI, O. & TANAKA, K. (1998). Mutational analysis of a function of xeroderma pigmentosum group A (XPA) protein in strand-specific DNA repair. *Nucleic Acids Research* **26**, 4662–4668.
- KOCH, S. C., KUPER, J., GASTEIGER, K. L., SIMON, N., STRASSER, R., EISEN, D., GEIGER, S., SCHNEIDER, S., KISKER, C. & CARELL, T. (2015). Structural insights into the recognition of cisplatin and AAF-dG lesion by Rad14 (XPA). *Proceedings of the National Academy of Sciences of the United States of America* **112**, 8272–8277.
- KONDRATICK, C. M., WASHINGTON, M. T., PRAKASH, S. & PRAKASH, L. (2001). Acidic residues critical for the activity and biological function of yeast DNA polymerase ϵ . *Molecular and Cellular Biology* **21**, 2018–2025.
- KUPER, J., BRAUN, C., ELIAS, A., MICHELS, G., SAUER, F., SCHMITT, D. R., POTERSZMAN, A., EGLY, J. M. & KISKER, C. (2014). In TFIIH, XPD helicase is exclusively devoted to DNA repair. *PLoS Biology* **12**, e1001954.
- KUPER, J., WOLSKI, S. C., MICHELS, G. & KISKER, C. (2012). Functional and structural studies of the nucleotide excision repair helicase XPD suggest a polarity for DNA translocation. *The EMBO Journal* **31**, 494–502.
- KURAOKA, I., MORITA, E. H., SAJO, M., MATSUDA, T., MORIKAWA, K., SHIRAKAWA, M. & TANAKA, K. (1996). Identification of a damaged-DNA binding domain of the XPA protein. *Mutation Research* **362**, 87–95.
- KUSUMOTO, R., MASUTANI, C., SUGASAWA, K., IWAI, S., ARAKI, M., UCHIDA, A., MIZUKOSHI, T. & HANAOKA, F. (2001). Diversity of the damage recognition step in the global genomic nucleotide excision repair *in vitro*. *Mutation Research* **485**, 219–227.
- LEHMANN, A. R. (2001). The xeroderma pigmentosum group D (XPD) gene: one gene, two functions, three diseases. *Genes & Development* **15**, 15–23.

- LEHMANN, A. R. (2003). DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. *Biochimie* **85**, 1101–1111.
- LEHMANN, A. R. (2008). XPD structure reveals its secrets. *DNA Repair* **7**, 1912–1915.
- LEHMANN, A. R., KIRK-BELL, S., ARLETT, C. F., PATERSON, M. C., LOHMAN, P. H., DE WEERD-KASTELEIN, E. A. & BOOTSMA, D. (1975). Xeroderma pigmentosum cells with normal levels of excision repair have a defect in DNA synthesis after UV-irradiation. *Proceedings of the National Academy of Sciences of the United States of America* **72**, 219–223.
- LI, L., ELLEDGE, S. J., PETERSON, C. A., BALES, E. S. & LEGERSKI, R. J. (1994). Specific association between the human DNA repair proteins XPA and ERCC1. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 5012–5016.
- LI, C., GOLEBIOWSKI, F. M., ONISHI, Y., SAMARA, N. L., SUGASAWA, K. & YANG, W. (2015). Tripartite DNA lesion recognition and verification by XPC, TFIIH, and XPA in nucleotide excision repair. *Molecular Cell* **59**, 1025–1034.
- LI, L., LU, X., PETERSON, C. A. & LEGERSKI, R. J. (1995a). An interaction between the DNA repair factor XPA and replication protein A appears essential for nucleotide excision repair. *Molecular and Cellular Biology* **15**, 5396–5402.
- LI, L., PETERSON, C. A., LU, X. & LEGERSKI, R. J. (1995b). Mutations in XPA that prevent association with ERCC1 are defective in nucleotide excision repair. *Molecular and Cellular Biology* **15**, 1993–1998.
- LIEBER, M. R. (1997). The FEN-1 family of structure-specific nucleases in eukaryotic DNA replication, recombination and repair. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* **19**, 233–240.
- LIU, H., RUDOLF, J., JOHNSON, K. A., MCMAHON, S. A., OKE, M., CARTER, L., MCROBBIE, A. M., BROWN, S. E., NAISMITH, J. H. & WHITE, M. F. (2008). Structure of the DNA repair helicase XPD. *Cell* **133**, 801–812.
- LIU, Y., YANG, Z., UTZAT, C., WANG, G., BASU, A. K. & ZOU, Y. (2005). Cooperative interaction of human XPA stabilizes and enhances specific binding of XPA to DNA damage. *Biochemistry* **44**, 7361–7368.
- LIVNEH, Z. (2001). DNA damage control by novel DNA polymerases: translesion replication and mutagenesis. *The Journal of Biological Chemistry* **276**, 25639–25642.
- LUJSTERBURG, M. S., GOEDHART, J., MOSER, J., KOOL, H., GEVERTS, B., HOUTSMULLER, A. B., MULLENDERS, L. H., VERMEULEN, W. & VAN DRIEL, R. (2007). Dynamic *in vivo* interaction of DDB2 E3 ubiquitin ligase with UV-damaged DNA is independent of damage-recognition protein XPC. *Journal of Cell Science* **120**, 2706–2716.
- LUJSTERBURG, M. S., VON BORNSTAEDET, G., GOURDIN, A. M., POLITI, A. Z., MONE, M. J., WARMERDAM, D. O., GOEDHART, J., VERMEULEN, W., VAN DRIEL, R. & HOFER, T. (2010). Stochastic and reversible assembly of a multiprotein DNA repair complex ensures accurate target site recognition and efficient repair. *The Journal of Cell Biology* **189**, 445–463.
- LUKIN, M. & DE LOS SANTOS, C. (2006). NMR structures of damaged DNA. *Chemical Reviews* **106**, 607–686.
- MAILLARD, O., SOLYOM, S. & NAEGELI, H. (2007). An aromatic sensor with aversion to damaged strands confers versatility to DNA repair. *PLoS Biology* **5**(4), e79.
- MALTSEVA, E. A., RECHKUNOVA, N. I., GILLET, L. C., PETRUSEVA, I. O., SCHARER, O. D. & LAVRIK, O. I. (2007). Crosslinking of the NER damage recognition proteins XPC-HR23B, XPA and RPA to photoreactive probes that mimic DNA damages. *Biochimica et Biophysica Acta* **1770**, 781–789.
- MASUTANI, C., ARAKI, M., YAMADA, A., KUSUMOTO, R., NOGIMORI, T., MAEKAWA, T., IWAI, S. & HANAOKA, F. (1999a). Xeroderma pigmentosum variant (XP-V) correcting protein from HeLa cells has a thymine dimer bypass DNA polymerase activity. *The EMBO Journal* **18**, 3491–3501.
- MASUTANI, C., KUSUMOTO, R., IWAI, S. & HANAOKA, F. (2000). Mechanisms of accurate translesion synthesis by human DNA polymerase η . *The EMBO Journal* **19**, 3100–3109.
- MASUTANI, C., KUSUMOTO, R., YAMADA, A., DOHMAE, N., YOKOI, M., YUASA, M., ARAKI, M., IWAI, S., TAKIO, K. & HANAOKA, F. (1999b). The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase η . *Nature* **399**, 700–704.
- MATHIEU, N., KACZMAREK, N. & NAEGELI, H. (2010). Strand- and site-specific DNA lesion demarcation by the xeroderma pigmentosum group D helicase. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 17545–17550.
- MATHIEU, N., KACZMAREK, N., RUTHEMANN, P., LUCH, A. & NAEGELI, H. (2013). DNA quality control by a lesion sensor pocket of the xeroderma pigmentosum group D helicase subunit of TFIIH. *Current Biology* **23**, 204–212.
- MATSUDA, T., BEBENEK, K., MASUTANI, C., HANAOKA, F. & KUNKEL, T. A. (2000). Low fidelity DNA synthesis by human DNA polymerase- η . *Nature* **404**, 1011–1013.
- MELLON, I., SPIVAK, G. & HANAWALT, P. C. (1987). Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. *Cell* **51**, 241–249.
- MIETUS, M., NOWAK, E., JACIUK, M., KUSTOSZ, P., STUDNICKA, J. & NOWOTNY, M. (2014). Crystal structure of the catalytic core of Rad2: insights into the mechanism of substrate binding. *Nucleic Acids Research* **42**, 10762–10775.
- MIN, J. H. & PAVLETICH, N. P. (2007). Recognition of DNA damage by the Rad4 nucleotide excision repair protein. *Nature* **449**, 570–575.
- MISSURA, M., BUTERIN, T., HINDGES, R., HUBSCHER, U., KASPARKOVA, J., BRABEC, V. & NAEGELI, H. (2001). Double-check probing of DNA bending and unwinding by XPA-RPA: an architectural function in DNA repair. *The EMBO Journal* **20**, 3554–3564.
- MITCHELL, D. L. & NAIRN, R. S. (1989). The biology of the (6–4) photoproduct. *Photochemistry and Photobiology* **49**, 805–819.
- MIYAMOTO, I., MIURA, N., NIWA, H., MIYAZAKI, J. & TANAKA, K. (1992). Mutational analysis of the structure and function of the xeroderma pigmentosum group A complementing protein. Identification of essential domains for nuclear localization and DNA excision repair. *The Journal of Biological Chemistry* **267**, 12182–12187.
- MOGGS, J. G., SZYMKOWSKI, D. E., YAMADA, M., KARRAN, P. & WOOD, R. D. (1997). Differential human nucleotide excision repair of paired and mismatched cisplatin-DNA adducts. *Nucleic Acids Research* **25**, 480–491.

- MOGGS, J. G., YAREMA, K. J., ESSIGMANN, J. M. & WOOD, R. D. (1996). Analysis of incision sites produced by human cell extracts and purified proteins during nucleotide excision repair of a 1,3-intrastrand d(GpTpG)-cisplatin adduct. *The Journal of Biological Chemistry* **271**, 7177–7186.
- MONE, M. J., VOLKER, M., NIKAIIDO, O., MULLENDERS, L. H., VAN ZEELAND, A. A., VERSCHURE, P. J., MANDERS, E. M. & VAN DRIEL, R. (2001). Local UV-induced DNA damage in cell nuclei results in local transcription inhibition. *EMBO Reports* **2**, 1013–1017.
- MORIKAWA, K. & SHIRAKAWA, M. (2000). Three-dimensional structural views of damaged-DNA recognition: T4 endonuclease V, *E. coli* Vsr protein, and human nucleotide excision repair factor XPA. *Mutation Research* **460**, 257–275.
- MORITA, E. H., OHKUBO, T., KURAOKA, I., SHIRAKAWA, M., TANAKA, K. & MORIKAWA, K. (1996). Implications of the zinc-finger motif found in the DNA-binding domain of the human XPA protein. *Genes to Cells* **1**, 437–442.
- MOSER, J., KOOL, H., GIAKZIDIS, J., CALDECOTT, K., MULLENDERS, L. & FOUSTERI, M. I. (2001). Sealing of Chromosomal DNA Nicks during Nucleotide Excision Repair Requires XRCC1 and DNA Ligase III α in a cell-cycle-specific manner. *Molecular Cells* **27**, 311–323.
- MOSER, J., KOOL, H., GIAKZIDIS, I., CALDECOTT, K., MULLENDERS, L. H. & FOUSTERI, M. I. (2007). Sealing of chromosomal DNA nicks during nucleotide excision repair requires XRCC1 and DNA ligase III α in a cell-cycle-specific manner. *Molecular Cell* **27**(2), 311–323.
- MU, D., BERTRAND-BURGGRAB, E., HUANG, J. C., FUCHS, R. P., SANCAR, A. & FUCHS, B. P. (1994). Human and *E. coli* excinucleases are affected differently by the sequence context of acetylaminofluorene-guanine adduct. *Nucleic Acids Research* **22**, 4869–4871.
- MU, D., PARK, C. H., MATSUNAGA, T., HSU, D. S., REARDON, J. T. & SANCAR, A. (1995). H reconstitution of human DNA repair excision nuclease in a highly defined system. *Journal of Biological Chemistry* **270**, 2415–2418.
- NAEGELI, H., BARDWELL, L. & FRIEDBERG, E. C. (1992). The DNA helicase and adenosine triphosphatase activities of yeast Rad3 protein are inhibited by DNA damage. A potential mechanism for damage-specific recognition. *The Journal of Biological Chemistry* **267**, 392–398.
- NAEGELI, H., MODRICH, P. & FRIEDBERG, E. C. (1993). The DNA helicase activities of Rad3 protein of *Saccharomyces cerevisiae* and helicase II of *Escherichia coli* are differentially inhibited by covalent and noncovalent DNA modifications. *The Journal of Biological Chemistry* **268**, 10386–10392.
- NAEGELI, H. & SUGASAWA, K. (2011). The xeroderma pigmentosum pathway: decision tree analysis of DNA quality. *DNA Repair* **10**, 673–683.
- NEWMAN, M., MURRAY-RUST, J., LALLY, J., RUDOLF, J., FADDEN, A., KNOWLES, P. P., WHITE, M. F. & McDONALD, N. Q. (2005). Structure of an XPF endonuclease with and without DNA suggests a model for substrate recognition. *The EMBO Journal* **24**, 895–905.
- NICHOLS, A. F., ONG, P. & LINN, S. (1996). Mutations specific to the xeroderma pigmentosum group E Ddb-phenotype. *The Journal of Biological Chemistry* **271**, 24317–24320.
- NISHINO, T., KOMORI, K., ISHINO, Y. & MORIKAWA, K. (2003). X-ray and biochemical anatomy of an archaeal XPF/Rad1/Mus81 family nuclease: similarity between its endonuclease domain and restriction enzymes. *Structure* **11**, 445–457.
- NOCENTINI, S., COIN, F., SAJO, M., TANAKA, K. & EGLY, J. M. (1997). DNA damage recognition by XPA protein promotes efficient recruitment of transcription factor II H. *The Journal of Biological Chemistry* **272**, 22991–22994.
- O'DONOVAN, A., DAVIES, A. A., MOGGS, J. G., WEST, S. C. & WOOD, R. D. (1994a). XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair. *Nature* **371**, 432–435.
- O'DONOVAN, A., SCHERLY, D., CLARKSON, S. G. & WOOD, R. D. (1994b). Isolation of active recombinant XPG protein, a human DNA repair endonuclease. *The Journal of Biological Chemistry* **269**, 15965–15968.
- OGI, T., LIMSIRICHAIKUL, S., OVERMEER, R. M., VOLKER, M., TAKENAKA, K., CLONEY, R., NAKAZAWA, Y., NIIMI, A., MIKI, Y., JASPERS, N. G., MULLENDERS, L. H., YAMASHITA, S., FOUSTERI, M. I. & LEHMANN, A. R. (2010). Three DNA polymerases, recruited by different mechanisms, carry out NER repair synthesis in human cells. *Molecular Cell* **37**(5), 714–727.
- OHMORI, H., FRIEDBERG, E. C., FUCHS, R. P., GOODMAN, M. F., HANAOKA, F., HINKLE, D., KUNKEL, T. A., LAWRENCE, C. W., LIVNEH, Z., NOHMI, T., PRAKASH, L., PRAKASH, S., TODO, T., WALKER, G. C., WANG, Z. & WOODGATE, R. (2001). The Y-family of DNA polymerases. *Molecular Cell* **8**, 7–8.
- OKSENYCH, V., BERNARDES DE JESUS, B., ZHOVMER, A., EGLY, J. M. & COIN, F. (2009). Molecular insights into the recruitment of TFIIH to sites of DNA damage. *The EMBO Journal* **28**, 2971–2980.
- ORANS, J., MCSWEENEY, E. A., IYER, R. R., HAST, M. A., HELLINGA, H. W., MODRICH, P. & BEESE, L. S. (2011). Structures of human exonuclease 1 DNA complexes suggest a unified mechanism for nuclease family. *Cell* **145**, 212–223.
- PARK, C. H., BESSHO, T., MATSUNAGA, T. & SANCAR, A. (1995a). Purification and characterization of the XPF-ERCC1 complex of human DNA repair excision nuclease. *The Journal of Biological Chemistry* **270**, 22657–22660.
- PARK, C. H., MU, D., REARDON, J. T. & SANCAR, A. (1995b). The general transcription-repair factor TFIIH is recruited to the excision repair complex by the XPA protein independent of the TFIIIE transcription factor. *The Journal of Biological Chemistry* **270**, 4896–4902.
- PARK, C. H. & SANCAR, A. (1994). Formation of a ternary complex by human XPA, ERCC1, and ERCC4 (XPF) excision repair proteins. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 5017–5021.
- PUGH, R. A., WU, C. G. & SPIES, M. (2012). Regulation of translocation polarity by helicase domain 1 in SF2B helicases. *The EMBO Journal* **31**, 503–514.
- RADEMAKERS, S., VOLKER, M., HOOGBRATEN, D., NIGG, A. L., MONE, M. J., VAN ZEELAND, A. A., HOEIJMAKERS, J. H., HOUTSMULLER, A. B. & VERMEULEN, W. (2003). Xeroderma pigmentosum group A protein loads as a separate factor onto DNA lesions. *Molecular and Cellular Biology* **23**, 5755–5767.
- RAJSKI, S. R., JACKSON, B. A. & BARTON, J. K. (2000). DNA repair: models for damage and mismatch recognition. *Mutation Research* **447**, 49–72.

- RAOUL, S., BARDET, M. & CADET, J. (1995). Gamma irradiation of 2'-deoxyadenosine in oxygen-free aqueous solutions: identification and conformational features of formamidopyrimidine nucleoside derivatives. *Chemical Research in Toxicology* **8**, 924–933.
- READ, C. M., CARY, P. D., CRANE-ROBINSON, C., DRISCOLL, P. C. & NORMAN, D. G. (1993). Solution structure of a DNA-binding domain from HMGI. *Nucleic Acids Research* **21**, 3427–3436.
- REARDON, J. T., NICHOLS, A. F., KEENEY, S., SMITH, C. A., TAYLOR, J. S., LINN, S. & SANCAR, A. (1993). Comparative analysis of binding of human damaged DNA-binding protein (XPE) and *Escherichia coli* damage recognition protein (UvrA) to the major ultraviolet photoproducts: T[c, s]T, T[t, s]T, T[6–4]T, and T[Dewar]T. *The Journal of Biological Chemistry* **268**, 21301–21308.
- REARDON, J. T. & SANCAR, A. (2002). Molecular anatomy of the human excision nuclease assembled at sites of DNA damage. *Molecular and Cellular Biology* **22**, 5938–5945.
- REARDON, J. T. & SANCAR, A. (2003). Recognition and repair of the cyclobutane thymine dimer, a major cause of skin cancers, by the human excision nuclease. *Genes & Development* **17**, 2539–2551.
- REARDON, J. T. & SANCAR, A. (2005). Nucleotide excision repair. *Progress in Nucleic Acid Research and Molecular Biology* **79**, 183–235.
- REISSNER, T., SCHNEIDER, S., SCHORR, S. & CARELL, T. (2010). Crystal structure of a cisplatin-(1,3-GTG) cross-link within DNA polymerase ϵ . *Angewandte Chemie International Edition* **49**, 3077–3080.
- RIEDL, T., HANAOKA, F. & EGLY, J. M. (2003). The comings and goings of nucleotide excision repair factors on damaged DNA. *The EMBO Journal* **22**, 5293–5303.
- ROBINS, P., JONES, C. J., BIGGERSTAFF, M., LINDAHL, T. & WOOD, R. D. (1991). Complementation of DNA repair in xeroderma pigmentosum group A cell extracts by a protein with affinity for damaged DNA. *The EMBO Journal* **10**, 3913–3921.
- RUPP, W. D. & HOWARD-FLANDERS, P. (1968). Discontinuities in the DNA synthesized in an excision-defective strain of *Escherichia coli* following ultraviolet irradiation. *Journal of Molecular Biology* **31**, 291–304.
- SAIJO, M., KURAOKA, I., MASUTANI, C., HANAOKA, F. & TANAKA, K. (1996). Sequential binding of DNA repair proteins RPA and ERCC1 to XPA *in vitro*. *Nucleic Acids Research* **24**, 4719–4724.
- SANCAR, A. & TANG, M. S. (1993). Nucleotide excision repair. *Photochemistry and Photobiology* **57**, 905–921.
- SCHAEFFER, L., ROY, R., HUMBERT, S., MONCOLLIN, V., VERMEULEN, W., HOEIJMAKERS, J. H., CHAMBON, P. & EGLY, J. M. (1993). DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. *Science* **260**, 58–63.
- SCHARER, O. D. (2013). Nucleotide excision repair in eukaryotes. *Cold Spring Harbor Perspectives Biology* **5**, a012609.
- SCHORR, S. & CARELL, T. (2010a). Mechanism of acetylaminofluorene-dG induced frameshifting by polymerase ϵ . *ChemBiochem* **11**, 2534–2537.
- SCHORR, S., SCHNEIDER, S., LAMMENS, K., HOPFNER, K. P. & CARELL, T. (2010b). Mechanism of replication blocking and bypass of Y-family polymerase η by bulky acetylaminofluorene DNA adducts. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 20720–20725.
- SCHWEIZER, U., HEY, T., LIPPS, G. & KRAUSS, G. (1999). Photocrosslinking locates a binding site for the large subunit of human replication protein A to the damaged strand of cisplatin-modified DNA. *Nucleic Acids Research* **27**, 3183–3189.
- SCRIMA, A., KONICKOVA, R., CZYZEWSKI, B. K., KAWASAKI, Y., JEFFREY, P. D., GROISMAN, R., NAKATANI, Y., IWAI, S., PAVLETICH, N. P. & THOMA, N. H. (2008). Structural basis of UV DNA-damage recognition by the DDB1–DDB2 complex. *Cell* **135**, 1213–1223.
- SHIN, D. S., PELLEGRINI, L., DANIELS, D. S., YELENT, B., CRAIG, L., BATES, D., YU, D. S., SHIVJI, M. K., HITOMI, C., ARVAI, A. S., VOLKMAN, N., TSURUTA, H., BLUNDELL, T. L., VENKITARAMAN, A. R. & TAINER, J. A. (2003). Full-length archaeal Rad51 structure and mutants: mechanisms for RAD51 assembly and control by BRCA2. *The EMBO Journal* **22**, 4566–4576.
- SHIVJI, M. K., PODUST, V. N., HUBSCHER, U. & WOOD, R. D. (1995). Nucleotide excision repair DNA synthesis by DNA polymerase epsilon in the presence of PCNA, RFC, and RPA. *Biochemistry* **34**, 5011–5017.
- SIJBERS, A. M., DE LAAT, W. L., ARIZA, R. R., BIGGERSTAFF, M., WEI, Y. F., MOGGS, J. G., CARTER, K. C., SHELL, B. K., EVANS, E., DE JONG, M. C., RADEMAKERS, S., DE ROOIJ, J., JASPERS, N. G., HOEIJMAKERS, J. H. & WOOD, R. D. (1996). Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell* **86**, 811–822.
- SILVERSTEIN, T. D., JAIN, R., JOHNSON, R. E., PRAKASH, L., PRAKASH, S. & AGGARWAL, A. K. (2010a). Structural basis for error-free replication of oxidatively damaged DNA by yeast DNA polymerase ϵ . *Structure* **18**, 1463–1470.
- SILVERSTEIN, T. D., JOHNSON, R. E., JAIN, R., PRAKASH, L., PRAKASH, S. & AGGARWAL, A. K. (2010b). Structural basis for the suppression of skin cancers by DNA polymerase ϵ . *Nature* **465**, 1039–1043.
- SINGH, S., FOLKERS, G. E., BONVIN, A. M., BOELEN, R., WECHSELBERGER, R., NIZTAYEV, A. & KAPTEIN, R. (2002). Solution structure and DNA-binding properties of the C-terminal domain of UvrC from *E. coli*. *The EMBO Journal* **21**, 6257–6266.
- STARESINIC, L., FAGBEMI, A. F., ENZLIN, J. H., GOURDIN, A. M., WIJGERS, N., DUNAND-SAUTHIER, I., GIGLIA-MARI, G., CLARKSON, S. G., VERMEULEN, W. & SCHARER, O. D. (2009). Coordination of dual incision and repair synthesis in human nucleotide excision repair. *The EMBO Journal* **28**, 1111–1120.
- STORY, R. M. & STEITZ, T. A. (1992). Structure of the recA protein–ADP complex. *Nature* **355**, 374–376.
- SUGASAWA, K., NG, J. M., MASUTANI, C., IWAI, S., VAN DER SPEK, P. J., EKER, A. P., HANAOKA, F., BOOTSMA, D. & HOEIJMAKERS, J. H. (1998). Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. *Molecular Cell* **2**, 223–232.
- SUGASAWA, K., OKAMOTO, T., SHIMIZU, Y., MASUTANI, C., IWAI, S. & HANAOKA, F. (2001). A multistep damage recognition mechanism for global genomic nucleotide excision repair. *Genes & Development* **15**, 507–521.
- SUGASAWA, K., OKUDA, Y., SAIJO, M., NISHI, R., MATSUDA, N., CHU, G., MORI, T., IWAI, S., TANAKA, K., TANAKA, K., HANAOKA, F. (2005). UV-induced ubiquitylation of XPC protein mediated by UV-DDB-ubiquitin ligase complex. *Cell* **121**, 387–400.
- SUGASAWA, K., SHIMIZU, Y., IWAI, S. & HANAOKA, F. (2002). A molecular mechanism for DNA damage recognition by the xeroderma pigmentosum group C protein complex. *DNA Repair* **1**, 95–107.

- SUGITANI, N., SHELL, S. M., SOSS, S. E. & CHAZIN, W. J. (2014). Redefining the DNA-binding domain of human XPA. *Journal of the American Chemical Society* **136**, 10830–10833.
- SUNG, P., HIGGINS, D., PRAKASH, L. & PRAKASH, S. (1988). Mutation of lysine-48 to arginine in the yeast RAD3 protein abolishes its ATPase and DNA helicase activities but not the ability to bind ATP. *The EMBO Journal* **7**, 3263–3269.
- SWEDER, K. S. & HANAWALT, P. C. (1992). Preferential repair of cyclobutane pyrimidine dimers in the transcribed strand of a gene in yeast chromosomes and plasmids is dependent on transcription. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 10696–10700.
- TAKAO, M., ABRAMIC, M., MOOS, M. JR., OTRIN, V. R., WOOTTON, J. C., MCLENIGAN, M., LEVINE, A. S. & PROTIC, M. (1993). A 127 kDa component of a UV-damaged DNA-binding complex, which is defective in some xeroderma pigmentosum group E patients, is homologous to a slime mold protein. *Nucleic Acids Research* **21**, 4111–4118.
- TANG, J. Y., HWANG, B. J., FORD, J. M., HANAWALT, P. C. & CHU, G. (2000). Xeroderma pigmentosum p48 gene enhances global genomic repair and suppresses UV-induced mutagenesis. *Molecular Cell* **5**, 737–744.
- TANTIN, D. (1998). RNA polymerase II elongation complexes containing the Cockayne syndrome group B protein interact with a molecular complex containing the transcription factor IIH components xeroderma pigmentosum B and p62. *The Journal of Biological Chemistry* **273**, 27794–27799.
- TANTIN, D., KANSAL, A. & CAREY, M. (1997). Recruitment of the putative transcription-repair coupling factor CSB/ERCC6 to RNA polymerase II elongation complexes. *Molecular and Cellular Biology* **17**, 6803–6814.
- TAPIAS, A., AURIOL, J., FORGET, D., ENZLIN, J. H., SCHARER, O. D., COIN, F., COULOMBE, B. & EGLY, J. M. (2004). Ordered conformational changes in damaged DNA induced by nucleotide excision repair factors. *The Journal of Biological Chemistry* **279**, 19074–19083.
- THOMA, B. S. & VASQUEZ, K. M. (2003). Critical DNA damage recognition functions of XPC-hHR23B and XPA-RPA in nucleotide excision repair. *Molecular Carcinogenesis* **38**, 1–13.
- THOREL, F., CONSTANTINOU, A., DUNAND-SAUTHER, I., NOUSPIKEL, T., LALLE, P., RAAMS, A., JASPERS, N. G., VERMEULEN, W., SHIVJI, M. K., WOOD, R. D. & CLARKSON, S. G. (2004). Definition of a short region of XPG necessary for TFIIH interaction and stable recruitment to sites of UV damage. *Molecular and Cellular Biology* **24**, 10670–10680.
- TIRODE, F., BUSSO, D., COIN, F. & EGLY, J. M. (1999). Reconstitution of the transcription factor TFIIH: assignment of functions for the three enzymatic subunits, XPB, XPD, and cdk7. *Molecular Cell* **3**, 87–95.
- TORNALETTI, S. & HANAWALT, P. C. (1999). Effect of DNA lesions on transcription elongation. *Biochimie* **81**, 139–146.
- TRINCAO, J., JOHNSON, R. E., ESCALANTE, C. R., PRAKASH, S., PRAKASH, L. & AGGARWAL, A. K. (2001). Structure of the catalytic core of *S. cerevisiae* DNA polymerase ϵ : implications for translesion DNA synthesis. *Molecular Cell* **8**, 417–426.
- TRIPSANES, K., FOLKERS, G., AB, E., DAS, D., ODJIK, H., JASPERS, N. G., HOEIJMAKERS, J. H., KAPTEIN, R. & BOELEN, R. (2005). The structure of the human ERCC1/XPF interaction domains reveals a complementary role for the two proteins in nucleotide excision repair. *Structure* **13**, 1849–1858.
- TRIPSANES, K., FOLKERS, G. E., ZHENG, C., DAS, D., GRINSTEAD, J. S., KAPTEIN, R. & BOELEN, R. (2007). Analysis of the XPA and ssDNA-binding surfaces on the central domain of human ERCC1 reveals evidence for subfunctionalization. *Nucleic Acids Research* **35**, 5789–5798.
- TSODIKOV, O. V., ENZLIN, J. H., SCHARER, O. D. & ELLENBERGER, T. (2005). Crystal structure and DNA binding functions of ERCC1, a subunit of the DNA structure-specific endonuclease XPF-ERCC1. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 11236–11241.
- TSODIKOV, O. V., IVANOV, D., ORELLI, B., STARESINCIC, L., SHOSHANI, I., OBERMAN, R., SCHARER, O. D., WAGNER, G. & ELLENBERGER, T. (2007). Structural basis for the recruitment of ERCC1-XPF to nucleotide excision repair complexes by XPA. *The EMBO Journal* **26**, 4768–4776.
- TSUTAKAWA, S. E., CLASSEN, S., CHAPADOS, B. R., ARVAI, A. S., FINGER, L. D., GUENTHER, G., TOMLINSON, C. G., THOMPSON, P., SARKER, A. H., SHEN, B., COOPER, P. K., GRASBY, J. A. & TAINER, J. A. (2011). Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. *Cell* **145**, 198–211.
- TUTEJA, N. & TUTEJA, R. (1996). DNA helicases: the long unwinding road. *Nature Genetics* **13**, 11–12.
- VAN DER SPEK, P. J., EKER, A., RADEMAKERS, S., VISSER, C., SUGASAWA, K., MASUTANI, C., HANAOKA, F., BOOTSMA, D. & HOEIJMAKERS, J. H. (1996). XPC and human homologs of RAD23: intracellular localization and relationship to other nucleotide excision repair complexes. *Nucleic Acids Research* **24**, 2551–2559.
- VENEMA, J., VAN HOFFEN, A., KARCAGI, V., NATARAJAN, A. T., VAN ZEELAND, A. A. & MULLENDERS, L. H. (1991). Xeroderma pigmentosum complementation group C cells remove pyrimidine dimers selectively from the transcribed strand of active genes. *Molecular and Cellular Biology* **11**, 4128–4134.
- VOLKER, M., MONE, M. J., KARMAKAR, P., VAN HOFFEN, A., SCHUL, W., VERMEULEN, W., HOEIJMAKERS, J. H., VAN DRIEL, R., VAN ZEELAND, A. A. & MULLENDERS, L. H. (2001). Sequential assembly of the nucleotide excision repair factors *in vivo*. *Molecular Cell* **8**, 213–224.
- WAKASUGI, M., KAWASHIMA, A., MORIOKA, H., LINN, S., SANCAR, A., MORI, T., NIKAI, O. & MATSUNAGA, T. (2002). DDB accumulates at DNA damage sites immediately after UV irradiation and directly stimulates nucleotide excision repair. *The Journal of Biological Chemistry* **277**, 1637–1640.
- WAKASUGI, M. & SANCAR, A. (1998). Assembly, subunit composition, and footprint of human DNA repair excision nuclease. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 6669–6674.
- WAKASUGI, M. & SANCAR, A. (1999). Order of assembly of human DNA repair excision nuclease. *The Journal of Biological Chemistry* **274**, 18759–18768.
- WANG, M., MAHRENHOLZ, A. & LEE, S. H. (2000). RPA stabilizes the XPA-damaged DNA complex through protein–protein interaction. *Biochemistry* **39**, 6433–6439.



- WANG, H., ZHAI, L., XU, J., JOO, H. Y., JACKSON, S., ERDJUMENT-BROMAGE, H., TEMPST, P., XIONG, Y. & ZHANG, Y. (2006). Histone H3 and H4 ubiquitylation by the CUL4-DDB-ROC1 ubiquitin ligase facilitates cellular response to DNA damage. *Mol. Cell* **22**, 383–394.
- WEIR, H. M., KRAULIS, P. J., HILL, C. S., RAINE, A. R., LAUE, E. D. & THOMAS, J. O. (1993). Structure of the HMG box motif in the B-domain of HMG1. *The EMBO Journal* **12**, 1311–1319.
- WHITE, M. F. (2009). Structure, function and evolution of the XPD family of iron-sulfur-containing 5'→3' DNA helicases. *Biochemical Society Transactions* **37**, 547–551.
- WINKLER, G. S., ARAUJO, S. J., FIEDLER, U., VERMEULEN, W., COIN, F., EGLY, J. M., HOEIJMAKERS, J. H., WOOD, R. D., TIMMERS, H. T. & WEEDA, G. (2000). TFIIF with inactive XPD helicase functions in transcription initiation but is defective in DNA repair. *The Journal of Biological Chemistry* **275**, 4258–4266.
- WOLSKI, S. C., KUPER, J., HANZELMANN, P., TRUGLIO, J. J., CROTEAU, D. L., VAN HOUTEN, B. & KISKER, C. (2008). Crystal structure of the FeS cluster-containing nucleotide excision repair helicase XPD. *PLoS Biology* **6**, e149.
- WOOD, R. D. (1997). Nucleotide excision repair in mammalian cells. *The Journal of Biological Chemistry* **272**, 23465–23468.
- YAN, H., YUAN, W., VELCULESCU, V. E., VOGELSTEIN, B. & KINZLER, K. W. (2002). Allelic variation in human gene expression. *Science* **297**, 1143.
- YAN, S., WU, M., BUTERIN, T., NAEGELI, H., GEACINTOV, N. E. & BRODY, S. (2003). Role of base sequence context in conformational equilibria and nucleotide excision repair of benzo[a]pyrene diol epoxide–adenine adducts†. *Biochemistry* **42**, 2339–2354.
- YANG, Z., ROGINSKAYA, M., COLIS, L. C., BASU, A. K., SHELL, S. M., LIU, Y., MUSICH, P. R., HARRIS, C. M., HARRIS, T. M. & ZOU, Y. (2006). Specific and efficient binding of xeroderma pigmentosum complementation group A to double-strand/single-strand DNA junctions with 3'- and/or 5'-ssDNA branches. *Biochemistry* **45**, 15921–15930.
- YANG, Z. G., LIU, Y., MAO, L. Y., ZHANG, J. T. & ZOU, Y. (2002). Dimerization of human XPA and formation of XPA2-RPA protein complex. *Biochemistry* **41**, 13012–13020.
- YASUI, M., DONG, H., BONALA, R. R., SUZUKI, N., OHMORI, H., HANAOKA, F., JOHNSON, F., GROLLMAN, A. P. & SHIBUTANI, S. (2004). Mutagenic properties of 3-(deoxyguanosin-N2-yl)-2-acetylaminofluorene, a persistent acetylaminofluorene-derived DNA adduct in mammalian cells. *Biochemistry* **43**, 15005–15013.
- YEH, J. I., LEVINE, A. S., DU, S., CHINTE, U., GHODKE, H., WANG, H., SHI, H., HSIEH, C. L., CONWAY, J. F., VAN HOUTEN, B. & RAPIĆ-OTRIN, V. (2012). Damaged DNA induced UV-damaged DNA-binding protein (UV-DDB) dimerization and its roles in chromatinized DNA repair. *Proceedings of the National Academy of Sciences of the United States of America* **109**, E2737–E2746.

1.2. Bulky DNA-Addukte - Die Substrate der Nukleotidexzisionsreparatur

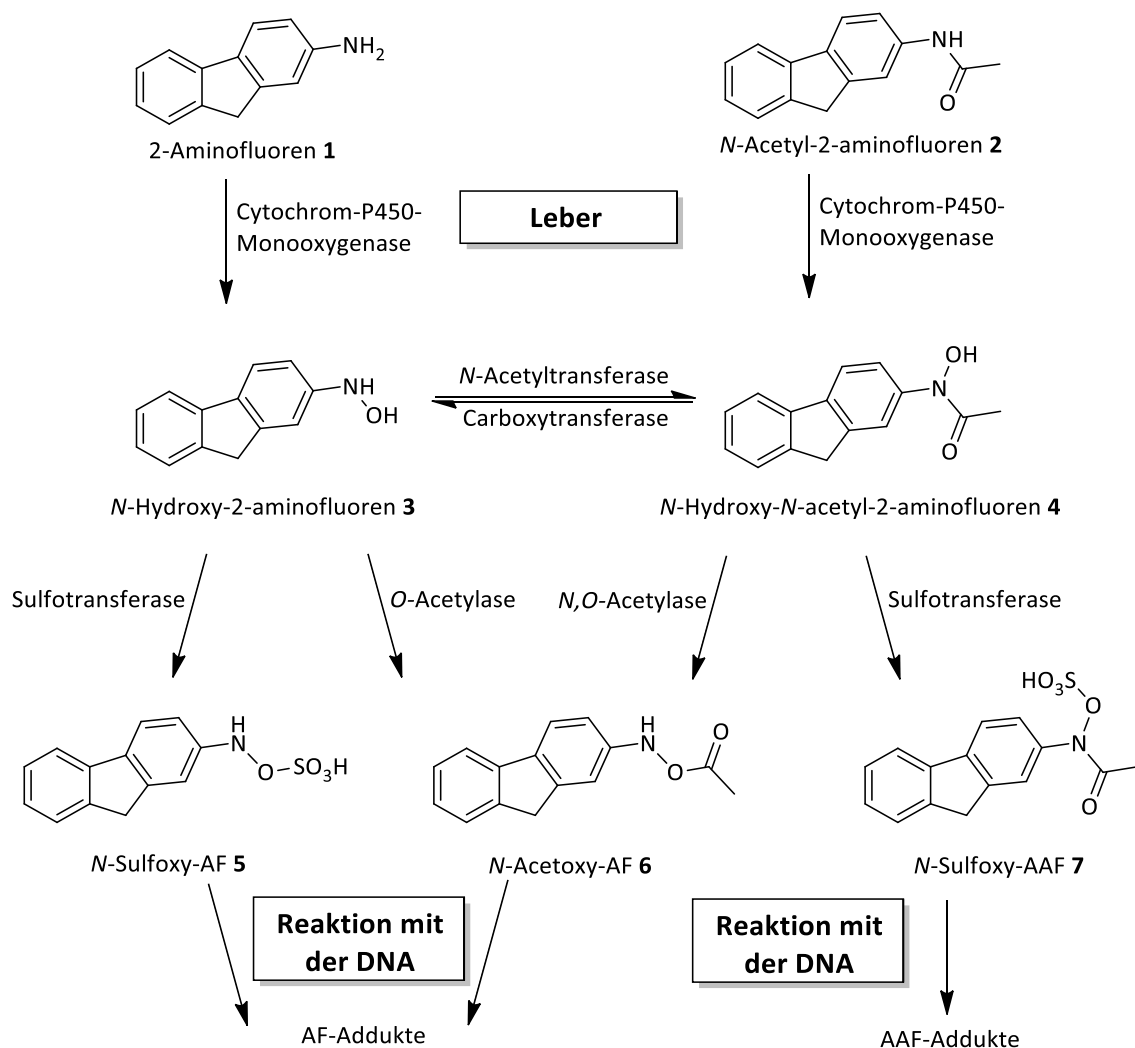
Die NER besitzt eine breite Substratspezifität und repariert unter anderem die sogenannten *bulky* DNA-Addukte. Solche DNA-Schäden können durch polyzyklische aromatische Kohlenwasserstoffe (PAKs), heterozyklische aromatische Amine, sowie aromatische Amine und Amide verursacht werden, indem diese durch Abbauprozesse im Körper in reaktive Metaboliten umgewandelt werden, um daraufhin mit der DNA reagieren zu können. Dabei entstehen die Addukte bevorzugt an der N^2 - oder an der C8-Position von Guanosin.^[1, 16-19]

Die PAKs entstehen unter anderem durch die unvollständige Verbrennung von fossilen Energieträgern und Kraftstoffen. Zudem können diese auch im Tabakrauch enthalten sein.^[18, 20-22] Infolge der krebserregenden Eigenschaften der PAKs und der weiten Verbreitung dieser Substanzen in Abgasen und Tabakrauch, ist ein deutliches Gesundheitsrisiko für den Menschen gegeben.^[18, 23-24]

Das gut untersuchte Benzo[*a*]pyren (B[*a*]P) gehört zu der PAKs Stoffklasse und kommt sowohl in Verbrennungsabgasen als auch in Tabakrauch vor.^[25] Dabei wird B[*a*]P bei Aufnahme in den Körper durch metabolischen Abbau, der normalerweise zur Entgiftung dient, zu einem starken Karzinogen umgewandelt. Der Metabolisierungsprozess erfolgt zuerst durch Epoxidierung und daraufhin durch Hydrolyse zu 1,2-Diolen. Das dabei entstehende, wasserlöslichere Diol kann vom Körper ausgeschieden werden. Dennoch kann ein restlicher Teil erneut durch Cytochrom-P450-Monooxygenasen epoxidiert werden, wodurch ein elektrophiles Epoxid entsteht, welches mit der N^2 -Position von Guanosin reagieren kann.^[22, 24-25]

Auch aromatische Amine sind indirekt wirkende Mutagene, die durch Metabolisierung aktiviert werden und daraufhin die DNA elektrophil angreifen. Sie kommen als Pyrolyseprodukte in Verbrennungsabgasen und Tabakrauch vor. Außerdem können diese auch beim Grillen und Braten von Fleisch bzw. Fisch durch die Maillard-Reaktion entstehen.^[20-21, 25-26]

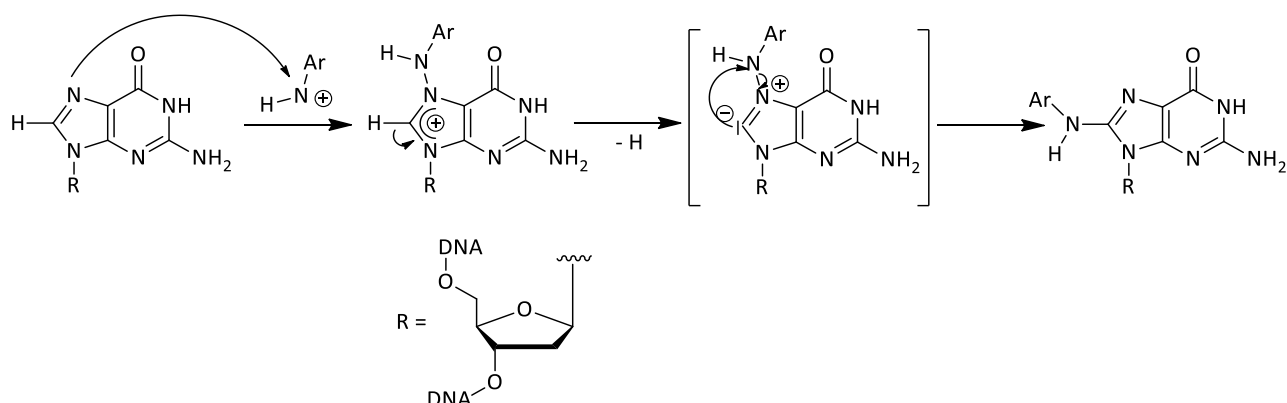
Als Modellverbindungen für diese Stoffklasse dient das 2-Aminofluoren (AF) und das *N*-Acetyl-2-aminofluoren (AAF), welche zur Erforschung der DNA-Reparatur verwendet werden. Das AF wurde ursprünglich als Insektizid entwickelt, aber aufgrund deren Kanzerogenität nie auf den Markt gebracht.^[27] Die metabolische Aktivierung von den Arylaminen AF (**1**) und AAF (**2**) ist in Schema 1 veranschaulicht.^[27-34]



Schema 1: Metabolische Aktivierung von den Arylaminen AF 1 und AAF 2. In Anlehnung nach *Grant et al.* und *Heflich et al.* ^[27, 29]

Die Metabolisierung der Arylamine dient zur Entgiftung des Körpers, wobei diese in der Leber aktiviert bzw. oxidiert werden. Dabei wird anfangs das aufgenommene Arylamin durch Cytochrom-P450-Monooxygenase in das entsprechende *N*-Hydroxy-2-aminofluoren 3 und *N*-Hydroxy-*N*-acetyl-2-aminofluoren 4 überführt. Zudem können die Derivate 3 und 4 durch eine *N*-Acetyltransferase oder einer Carboxyesterase ineinander umgewandelt werden. Weiterhin können diese Metaboliten in einer Entgiftungsreaktion durch die Glucuronosyltransferase zu dem Glucuronsäureester umgesetzt werden, welcher dann vom Körper ausgeschieden werden kann. Die Derivate 3 und 4 können mittels einer Sulfotransferase bzw. einer Acetyltransferase zu den reaktiven Estern 5, 6 und 7 umgewandelt werden. ^[27-35] Diese können über Intermediate, in denen sich ein elektrophiles Nitreniumion bildet, mit der DNA reagieren. Dabei entstehen meist C8- und *N*²-Addukte des 2'-Desoxyguanosins. ^[6] Auch die Metabolisierung anderer Arylamine, wie z.B. Anilin, 2-Aminonaphthalin, 1-Nitropyren und 2-Aminoanthracen, verläuft nach einem ähnlichen Schema wie

AAF und wirken kanzerogen. Denn diese werden auch durch Cytochrom-P450 aktiviert.^[19, 36-39] Anfangs wurde angenommen, dass das C8-Addukt durch Addition des Nitreniumions an die nukleophilere N7-Position des Guanosins entsteht. Daraufhin bildet sich ein intermediäres Ylid durch Deprotonierung an der C8-Position, woraufhin das Arylamin von der N7-Position zur C8-Position durch eine Stevens-ähnlichen-Umlagerung transferiert wird (Schema 2).^[40] Im Gegensatz dazu, postulieren *McClelland et al.* in späteren Studien, dass das Nitreniumion direkt an der C8-Position angreift.^[41]

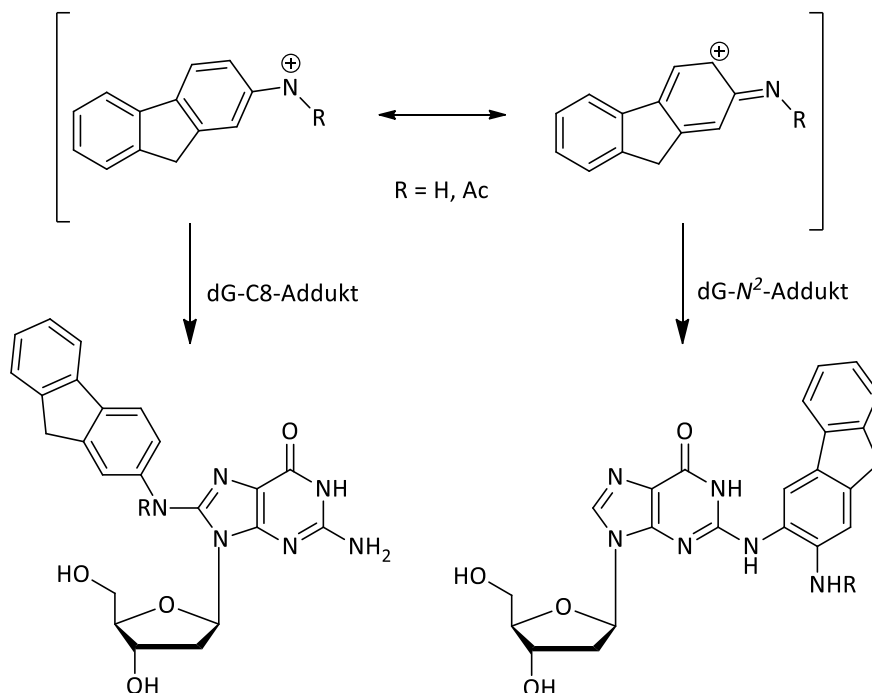


Schema 2: Postulierter Mechanismus zur Bildung eines C8-Adduktes durch ein Nitreniumion. In Anlehnung nach *Humphreys et al.*^[40]

Das C8-AF- und das C8-AAF-Guanosinaddukt haben, trotz Strukturähnlichkeit, unterschiedliche Wirkung auf die DNA und damit unterschiedliche Auswirkungen auf die NER. Das AAF nimmt eine *syn*-Konformation bezüglich der glykosidischen Bindung ein. Dabei kann das Fluoren in die DNA Doppelhelix interkalieren und das Guanosin wird in die große Furche verdrängt, wobei die Watson-Crick Basenpaarung nicht erhalten bleibt und der DNA-Duplex thermodynamisch destabilisiert wird.^[42-44] Durch die Interkalation des Fluorenteils im AAF und der sterisch anspruchsvollen Acetylgruppe kann es zu -1 und -2 *frameshift*-Mutationen kommen.^[28, 45] Im Gegensatz zum C8-AAF-Addukt liegt das C8-AF-Addukt in zwei Konformationen vor (*syn*- und *anti*-Konformation). Bei der *anti*-Konformation, welches die Hauptkonformation ist, liegt das Fluoren in der großen Furche, wobei der DNA-Duplex nicht beeinträchtigt wird und die Watson-Crick Basenpaarung erhalten bleibt.^[42-44] Dadurch kann die Polymerase den Schaden in der *anti*-Konformation überlesen, wobei das transstehende Fluoren zur Blockierung der Polymerase führt.^[46]

Durch *in vitro* Studien konnte gezeigt werden, dass bei der Schadensbildung 60% des C8-AAF-Addukts, 20% des C8-AF-Addukts und nur 5-15% des *N*²-AAF-Addukts entstehen. Im Vergleich dazu, konnte bei *in vivo* Studien in Leberzellen von Ratten belegt werden, dass größtenteils das C8-AF-Addukt (80%) und das C8- sowie das *N*²-AAF-Addukt (15% und 5%) nur in geringen Mengen

entstehen.^[47-49] Neben der Reaktion des Arylnitreniumions mit der C8-Position kann durch die Resonanzstabilisierung das Carbokation mit der N²-Position reagieren (Schema 3).^[50-51]



Schema 3: Mesomere Grenzstruktur von AF und AAF als Arylnitreniumion oder als Carbokation, welche mit 2'-Desoxyguanosin das C8- und N²-Addukt bilden können.^[50-51]

Zudem konnte in Studien von *Howard et al.* gezeigt werden, dass die C8-AF- und N²-AAF-Addukte in Tieren eine längere Lebensdauer haben als das C8-AAF-Addukt.^[48] Dabei wurde AAF männlichen Ratten injiziert, wobei das gebildete C8-AAF nur ein Tag nachweisbar war. Jedoch konnten die anderen beiden Addukte bis zu zwei Wochen später noch nachgewiesen werden. Daraus wurde geschlossen, dass das C8-AAF-Addukt im Körper unmittelbar erkannt und entfernt werden kann.^[47-48] Da das N²-AAF-Addukt in der kleinen Furche der DNA liegt und die DNA thermodynamisch stabilisiert, kann die Detektion des Schadens erschwert sein.^[52] *Kitsera et al.* gelang es in *knock-out* Experimenten die unterschiedlichen Reparaturraten von C8-AAF- und N²-AAF-dG zu erklären.^[53] Die C8-AAF-Addukte können leicht durch die globale Genomreparatur (GGR) erkannt und entfernt werden, wobei das N²-AAF-Addukt nur durch die Transkriptions-gekoppelte Reparatur (TCR), welche die Proteine CSA und CSB benötigt, entfernt werden kann.^[53] Weiterhin konnte gezeigt werden, dass XPA notwendig für die Reparatur beider Schäden ist, wobei XPC eine geringere Rolle spielt. Zudem wurden weitere *knock-out* Experimente während meiner Doktorarbeit durchgeführt, die zeigten, dass das C8-AF-Addukt durch die GGR entfernt wird. Dennoch werden 60% des Schadens bei der Transkription überlesen (siehe Abschnitt 3.4.).

1.3. Replikationsprotein A

1.3.1. Struktureller Aufbau von RPA

Das Replikation Protein A (RPA) ist ein wichtiges eukaryotisches, einzelsträngige (ss) DNA-bindendes Protein, welches eine wichtige Rolle in der Replikation, Reparatur und Rekombination spielt.^[54-57]

Der strukturelle Aufbau von RPA ist in Abbildung 3 veranschaulicht.^[58-60]

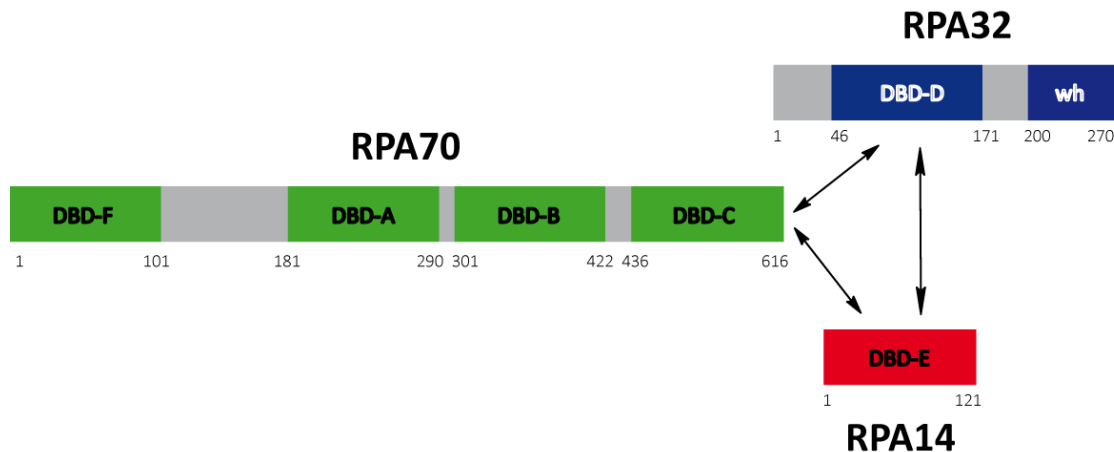


Abbildung 3: Schematische Darstellung der einzelnen RPA Domänen. Die Domänen der RPA70 Untereinheit sind in grün, RPA32 in blau und RPA14 in rot dargestellt. Der Trimerisierungskern besteht aus der DNA-bindenden Domäne C (DBD-C), DBD-D und -E. wh: *winged*-Helix Domäne.^[58-60]

RPA ist ein heterotrimeres Protein und wird durch drei verschiedene Gene kodiert, die jeweils für eine Untereinheit kodieren. Die drei Untereinheiten von RPA sind RPA70 (70 kDa), RPA32 (32 kDa) und RPA14 (14 kDa).^[54, 59, 61-62] Jede von ihnen hat eine oder mehr DNA-bindende-Domänen (DBD), welche aus einem Oligonukleotid/Oligosaccharid-bindenden (OB) Motiv zusammengesetzt und durch flexible Linker verbunden sind.^[63-64] Die größte Untereinheit RPA70 besitzt insgesamt vier Domänen, die DBD-A, DBD-B, DBD-C und eine N-terminale Protein-Protein Interaktionsdomäne (DBD-F).^[58-59, 65] Außerdem besitzt die DBD-C ein Zink-bindendes Motiv, wobei ein Zinkion an vier Cysteine gebunden ist.^[66] Die mittlere Untereinheit RPA32 umfasst die DBD-D und eine C-terminale Protein-Protein Interaktionsdomäne, die sogenannte *winged*-Helix Domäne (wh). Die N-terminale Region von RPA32 ist sehr ungeordnet und kann mehrfach phosphoryliert werden.^[61] Die kleinste Untereinheit RPA14 besteht aus einer Domäne (DBD-E), wobei dessen genaue Funktion bis heute noch nicht vollständig geklärt ist. Das trimere RPA bildet ein sehr stabilen Komplex, indem der Trimerisierungskern durch die DBD-C bis DBD-E gebildet wird.^[56, 58] Die langen Protein-Linker zwischen den einzelnen Domänen sind sehr flexibel und können unterschiedliche Konformationen einnehmen.^[61, 64]

1.3.2. RPA-DNA Bindung

RPA bindet bevorzugt an Polypyrimidin-Sequenzen und bindet ssDNA mit einer bis zu 1000-fach höheren Affinität als doppelsträngige (ds) DNA, wobei die DBD-A und DBD-B die höchste Affinität für DNA aufzeigen.^[67-69] Des Weiteren wird die DNA-Affinität von RPA durch die Länge der DNA beeinflusst, da eine kürzere ssDNA zu einer geringen Bindungsaffinität von RPA führt.^[59, 67, 70-71] Durch die flexible Struktur von RPA wird angenommen, dass RPA in verschiedenen Konformationen an die DNA, abhängig von deren Länge, binden kann.^[57, 68, 72-74] Die DBD-A und DBD-B, die eine hohe DNA-Affinität besitzen, sowie die DBD-C und DBD-E, die eine geringere DNA-Affinität aufweisen, ermöglichen die Bildung eines stabilen RPA/DNA-Komplexes. Dies konnte unter anderem durch Mutationsstudien belegt werden.^[57, 68, 75] Auch konnte kürzlich gezeigt werden, dass RPA dynamisch, durch die strukturelle Flexibilität, mit ssDNA interagieren und zügig entlang von ssDNA diffundieren kann.^[57, 72] Dabei kann RPA DNA-Haarnadelstrukturen oder sogar doppelsträngige Regionen aufschmelzen.^[57, 72] Wie bereits oben erwähnt, kann RPA in verschiedenen Konformationen an die DNA binden, wobei jede DBD wahrscheinlich mit vier bis sechs Nukleotiden (nt) der ssDNA interagiert. Dadurch kann ein instabiler 8nt Bindungsmodus, der durch *crosslinking*-Studien belegt wurde, oder ein stabiler 30nt Bindungsmodus entstehen. Der 8-10nt Bindungsmodus konnte auch beobachtet werden, wenn RPA mit kurzer DNA interagiert.^[67, 76-78] *Nguyen et al.* konnten durch thermodynamische Studien nur einen 18-20nt und einen 28-30nt Bindungsmodus von RPA identifizieren, woran drei bis vier DBD beteiligt sind.^[56, 72, 79] *Chen et al.* belegten 2016 unter anderem durch *Brown'sche* Dynamik Simulationen, dass der RPA-DNA Komplex in zwei kinetischen Zuständen vorliegt, und dass der stabilere Zustand, der durch schrittweise Interaktion der einzelnen RPA Domänen entsteht, für die DNA Reparatur benötigt.^[57, 80] Folglich stellen *Chen et al.* die Hypothese auf, dass zuerst eine bis zwei DBD (DBD-A und DBD-B) mit der DNA interagieren. Anschließend kann ein etwas stabilerer RPA/DNA-Komplex gebildet werden, indem drei bis vier DBD mit der DNA interagieren. Durch eine folgende Konformationsänderung des Komplexes oder durch Assoziation einer DBD, kann ein stabilerer Komplex gebildet werden.^[57] Diese Hypothese wird in der folgenden Abbildung 4 dargestellt.^[56-57]

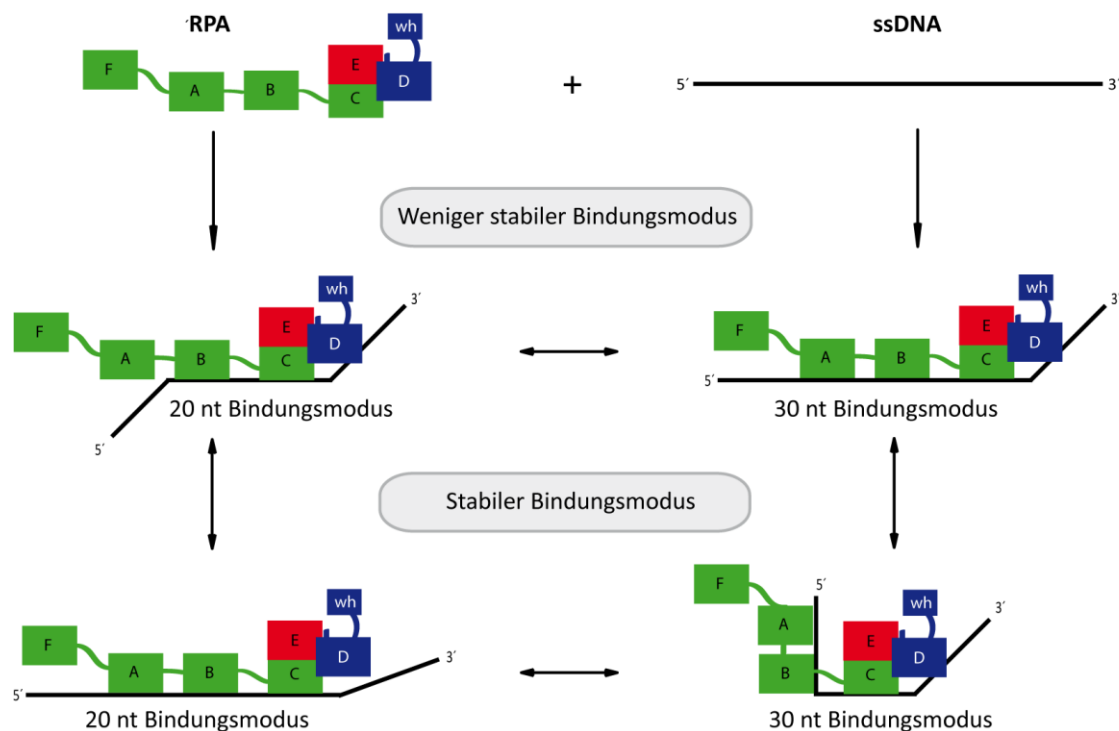


Abbildung 4: Schematische Darstellung von RPA in seinen verschiedenen, dynamischen Bindungsmodi. Die Domänen der RPA70 Untereinheit sind in grün, RPA32 in blau und RPA14 in rot dargestellt. RPA besitzt einen 20nt und einen 30nt Bindungsmodus. Durch dynamische Umlagerung kann RPA verschiedene DNA-Bindungskonformationen einnehmen. Die Abbildung wurde von *Chen und Wold* sowie von *Chen et al.* adaptiert.^[56-57]

1.3.3. Die verschiedenen Funktionen von RPA

RPA spielt eine wichtige Rolle in verschiedenen metabolischen Prozessen wie in der Replikation, Reparatur und Rekombination, um die Genomstabilität zu bewahren. In erster Linie entfernt RPA sekundäre Strukturen, wie z.B. Haarnadelstrukturen, schützt die ssDNA vor Nukleasen oder kann mit verschiedenen Proteinen interagieren.^[60, 81-82] Protein-Protein Interaktionen finden nicht nur durch die DBD-A bis DBD-D, sondern hauptsächlich durch die DBD-F oder durch die wh-Domäne statt. Eine Dysfunktion von RPA kann oft tödlich sein. Des Weiteren können Mutationen von RPA zu einer Fehlfunktion der DNA Reparatur oder zur Genominstabilität führen.^[56, 82-84] In Mutationsstudien von RPA konnte belegt werden, dass die maximale Bindungsaffinität zu ssDNA nicht mit der Funktion selbst korreliert. Unter anderem zeigten Mutationen in den DBD von RPA, die zu einer geringeren DNA-Bindungsaffinität führten, dennoch volle Funktion. Auch konnte gezeigt werden, dass eine hohe Affinität zu ssDNA nicht ausreichend für eine vollständige Funktion von RPA ist. Denn Mutationen in aromatischen Resten von der DBD-A oder DBD-B führten zu einem Funktionsverlust der DNA Reparatur, wohingegen die DNA Replikation nicht beeinträchtigt wurde. Folglich benötigt die DNA Reparatur und die Replikation unterschiedliche RPA-DNA Interaktionen.^[56, 83, 85]

Des Weiteren kann die Konformation und die Dynamik von RPA mit ssDNA durch Protein Interaktionen und posttranslationale Modifikationen von RPA variieren. RPA32 wird vor allem am N-Terminus phosphoryliert werden, wodurch Protein und DNA-Interaktionen reguliert werden können. Phosphorylierungen von RPA spielen eine Rolle in der homologen Rekombination und in der DNA Reparatur. Weiterhin können Lysine am C-Terminus von RPA70 sumoyliert werden, wenn DNA Doppelstrangbrüche vorliegen, was zur Rekrutierung von RAD51 an die DNA führt. Allerdings ist die genaue Regulierung durch Phosphorylierungen bzw. Sumoylierungen noch nicht ganz verstanden.^[56, 86-89]

Außerdem hat RPA eine wichtige Bedeutung für die Initiation und Elongation der S-Phase der DNA Replikation. Dabei kann RPA nicht nur die entstehende ssDNA am Replikationsursprung stabilisieren, sondern unterstützt die Rekrutierung von anderen Proteinen an die Replikationsgabel. Auch fördert RPA die Bindung der DNA-Polymerase an den Replikations-Initiationskomplex, indem es mit dieser interagiert.^[57, 59, 90-95] *Toledo et al.* belegten, dass RPA bei Replikationsstress enorm wichtig ist, da dabei die ssDNA Regionen stark ansteigen.^[82] Denn RPA stabilisiert die gebildeten Replikationsgabeln und fördert die ATR Aktivierung, um die Replikation zu stoppen. Dadurch können keine weiteren ssDNA Regionen gebildet werden. Durch *high-throughput* Mikroskopie konnten sie auch zeigen, dass eine zu hohe Anzahl an ssDNA Regionen und eine reduzierte Konzentration von RPA zu Doppelstrangbrüchen und diese wiederum zum Zelltod führen können.^[56, 82] Neueste Studien belegen, dass RPA auch eine Funktion beim Zusammenbau der Nukleosomen des Chromatins hat, indem es als eine Art Bindungsplattform für die Koordination von Histonen dient. Dabei kann RPA die Histone H3/H4 binden und fördert dadurch die Bindung der Histone an benachbarte dsDNA der Replikationsgabel. Folglich kann RPA den Nukleosom Zusammenbau und die laufende DNA Replikation koppeln.^[96]

Neben der Replikation, ist RPA auch in den verschiedenen DNA-Reparatur Mechanismen relevant, wie in der NER, BER, MMR oder Doppelstrangbruch (DSB) Reparatur. In der NER kann RPA mit verschiedenen Proteinen interagieren, wie z.B. XPA um den offenen DNA-Komplex zu stabilisieren. Die Interaktion mit XPA wird im Kapitel 1.5.1 genauer beschrieben. Zudem kann RPA die Nuklease XPG für die duale Inzision richtig positionieren. Außerdem ist RPA beim Schließen der Lücke im DNA-Doppelstrang mit dem *Proliferating-Cell-Nuclear-Antigen* (PCNA), Replikationsfaktor C (RFC) und der DNA Polymerase involviert.^[97-98] Die genaue Rolle von RPA in der BER ist noch nicht eindeutig verstanden. Bisher konnte nur die Interaktion von Uracil-DNA Glykosylase mit RPA identifiziert werden.^[99] In der MMR übernimmt RPA gleich mehrere Funktionen, wobei das Zinkfinger Motiv von RPA70 involviert ist. Unter anderem kann es die entstehenden ssDNA Regionen stabilisieren und fördert die DNA-Synthese während der Reparatur.^[59, 100] Die dynamischen Eigenschaften von RPA mit

ssDNA helfen bei der Reparatur von DNA-Doppelstrangbrüchen (DSB), die z.B. durch ionisierende Strahlung oder endogene bzw. exogene DNA-schädlichen Reagenzien entstehen können. Diese können unter anderem durch die homologe Rekombination (HR) repariert werden. Die Regulation von DSD durch RPA in der HR ist in der Abbildung 5 veranschaulicht.^[56, 59-60]

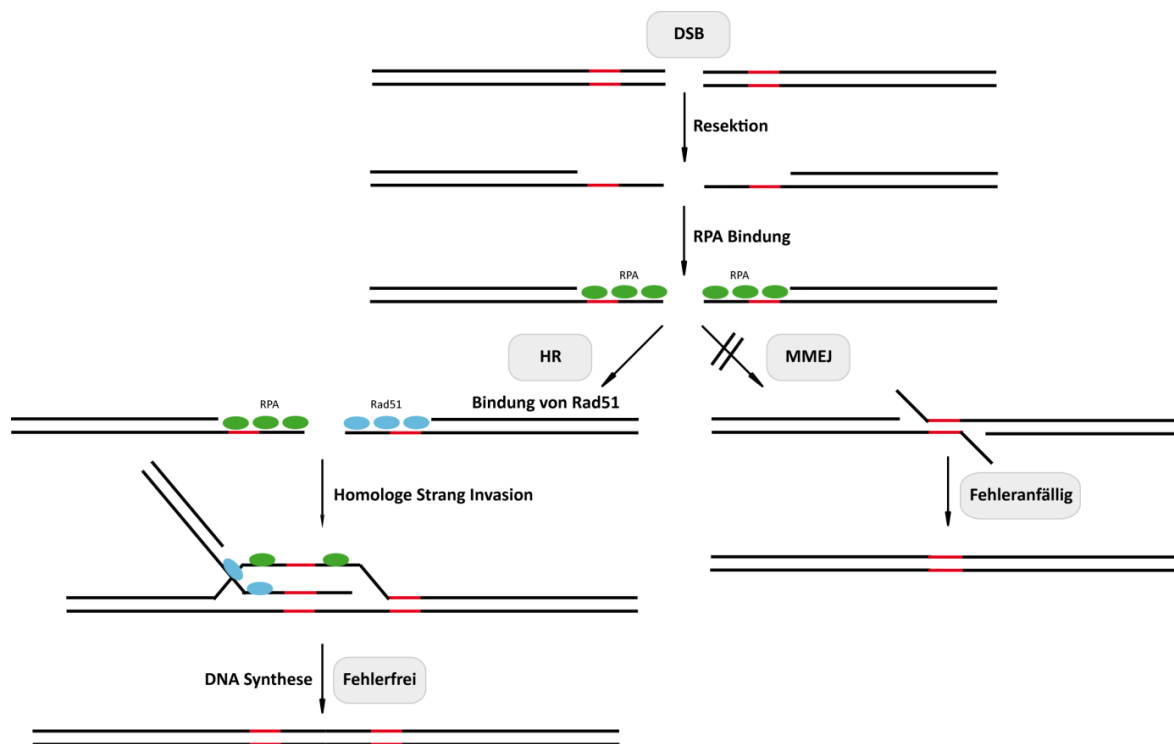


Abbildung 5: RPA bindet den ssDNA Überhang, der bei der Resektion von DSB entsteht. Durch die Stabilisierung des Überhangs inhibiert RPA die MMEJ und fördert die HR. Auch fördert RPA die Bindung von Rad51, welches RPA verdrängt und die HR katalysiert. RPA ist in grün, Rad51 in blau und die mikrohomologen Sequenzen sind in rot dargestellt. Die Abbildung wurde von Chen und Wold adaptiert.^[56]

In der HR wird durch nukleolytischen Abbau ein ca. 50-100nt ssDNA 3'-Überhang produziert, woraufhin RPA an den Überhang bindet und diesen stabilisiert. RPA fördert die HR durch Interaktion mit HR-Proteinen und Rad51, welches die Rekombination katalysiert. Rad51 kann durch die Mediatorproteine Rad52 und BRCA2 RPA von der ssDNA verdrängen, woraufhin Rad51 die Hybridisierung des Überhangs mit dem homologen DNA-Duplex katalysiert. Dies führt zur Ausbildung einer D-Schleife und die DNA-Synthese kann stattfinden. RPA kann bei der Strang Invasion den verdrängten Strang während der Rekombination stabilisieren. Des Weiteren kann der DSB auch durch MMEJ (*microhomology-mediated end joining*), die direkt mit der HR konkurriert, repariert werden. Im Gegensatz zu der HR ist die MMEJ jedoch fehleranfällig und genetisches Material geht verloren. Ferner fördert RPA die HR und inhibiert die MMEJ, indem es die Hybridisierung von mikrohomologen Sequenzen verhindert. Punktmutationen in den DBD von RPA führen zu einer erhöhten MMEJ, da diese schneller von der DNA dissoziieren und somit der ssDNA Überhang nicht mehr stabilisieren werden kann.^[56, 60, 81, 101-105]

1.3.4. Regulation von RPA – Allosterie

Obwohl RPA viele Bindungspartner hat, ist die Bindungsoberfläche für diese limitiert und die Bindungsseiten müssten für die Interaktionspartner überlappen. Es wird vermutet, dass RPA durch einen allosterischen Mechanismus die Bindungsstellen für Interaktionspartner regulieren kann. Durch die Bindung eines Liganden ändert sich die Konformation von allosterisch geregelten Proteinen. Erste Hinweise auf einen allosterischen Mechanismus von RPA entdeckten *Jackson et al.*^[106] Sie stellten fest, dass in der homologen Rekombination die Affinität für die ssDNA Bindung von RPA um einen Faktor fünf ansteigt, indem der Mediator Rad52 an RPA32C bindet. Diese Interaktion führt zu einer erhöhten Bindung der DBD-D an die ssDNA.^[106-107] Eine weitere interessante allosterische Interaktion konnte von *Jiang et al.* beobachtet werden.^[108] Diese fanden heraus, dass das *origin binding motif* (OBD) der Helikase SV40 T Antigen (Tag) mit den DBD-A und -B von RPA70 (RPA70AB) während der DNA Replikation interagieren kann, wodurch die ssDNA Bindungsaffinität allosterisch reguliert wird. Denn RPA70AB bildet mit der Tag-Helikase und einer 8nt ssDNA einen ternären Komplex, wohingegen bei einer 30nt ssDNA die Tag-Helikase von RPA dissoziiert. Deswegen postulieren *Jiang et al.*, dass sobald die Tag-Helikase mehr dsDNA entwindet, RPA sich von der kompakten tertiären Struktur in einen verlängerten ssDNA Bindungsmodus allosterisch umformt, wodurch die Tag-Bindung geschwächt wird. Diese Ergebnisse zeigen, dass RPA während der DNA-Replikation, vermittelt durch die Tag-Helikase, in zwei Schritten an die DNA gebunden wird.^[107-108]

1.3.5. Studien zur Strukturaufklärung von RPA

RPA ist ein sehr dynamisches und flexibles Protein und somit schwer zu kristallisieren. Dennoch konnten einige Gruppen in den letzten Jahren NMR- oder Kristallstrukturen von unterschiedlichen RPA Untereinheiten erhalten und somit wichtige Funktionen und Interaktionen von RPA aufdecken. Trotzdem ist bis heute unklar, wie die vollständige Struktur von RPA mit und ohne DNA aussieht und wäre daher von großer Bedeutung. Insbesondere Strukturen mit Interaktionsproteinen, wie mit XPA, wären von hohem Interesse.

1.3.5.1. Elektronenmikroskopie von RPA

Treuner et al. konnten erste Strukturen von RPA mit ssDNA durch Elektronenmikroskopie mittels Negativfärbung erhalten.^[109] Diese zeigten, dass der RPA/DNA-Komplex, abhängig von der

Salzkonzentration, drei verschiedene Konformationen einnehmen kann. Die Kontur der ssDNA, an der RPA gebunden ist, erscheint zickzackartig und unregelmäßig. Bei einer geringen Salzkonzentration nimmt die DNA im Komplex eine ausgedehnte Konformation ein, die ähnlich zur Protein-freien DNA ist. Durch Erhöhung der Salzkonzentration mit Natriumchlorid nimmt die DNA im Nukleoprotein Komplex eine kompaktere Konformation ein. Eine noch größere Kompaktion des RPA/DNA-Komplexes konnte durch Erhöhung der Magnesiumchlorid-Konzentrationen erreicht werden. Zudem zeigten Elektronenmikroskopiebilder, dass RPA dsDNA teilweise entwinden kann. Im Gegensatz zu anderen Einzelstrang-bindenden Proteinen (SSB) wickelt sich die ssDNA nicht um RPA herum, sondern RPA verlängert sich entlang der DNA. Dennoch konnte durch die Negativfärbung keine Struktur von RPA gelöst werden.^[73, 109-110] Witosch *et al.* konnten im Jahr 2014 eine Kryoelektronenmikroskopiestruktur von RPA im Komplex mit einem Tim-Tipin-Komplex und ssDNA in einer Auflösung von 17 Å erhalten.^[111] Der Tim-Tipin-Komplex ist in der Koordination der DNA Replikation involviert und wird vermutlich durch RPA an die Replikationsgabel rekrutiert. In dieser Struktur nimmt RPA eine hufeisenförmige Form an.^[111]

1.3.5.2. NMR Strukturen von RPA

NMR-Spektroskopie ist eine wichtige Methode, welche die Charakterisierung von Proteinen in Lösung ermöglicht. Mer *et al.* konnten die Struktur des C-Terminus von RPA32 (Aminosäure 170-270) durch NMR-Spektroskopie lösen (Abbildung 6a).^[112] In dieser wird ersichtlich, dass die ersten 32 Reste des C-Terminus unstrukturiert sind und daher wahrscheinlich einen flexiblen Linker zur DBD-D bilden. Zudem zeigt diese eine kompakte globuläre α/β wh-Domäne auf, die durch ein rechtsgängiges dreifach-Helix Bündel und drei kurze anti-parallele β -Stränge gebildet wird. Diese wh-Domäne ist wichtig für Protein-Protein-Interaktionen, wie die mit XPA. Anschließend untersuchten Mer *et al.* die NMR-Struktur von RPA32C mit einem 16 Aminosäuren kurzen Peptid, (Uracil-DNA Glykosylase, UNG₇₃₋₈₈), um die Bindungssequenz von RPA für XPA und Rad52 zu bestimmen. RPA32C bindet an das Peptid, welches eine α -Helix ausbildet, vor allem durch seinen zweiten und dritten β -Strang. Die Bindung wird hauptsächlich durch elektrostatische Interaktionen vermittelt, indem fünf basische Reste des Peptids mit zwei sauren Resten von RPA32C interagieren. Danach wurden kurze Peptide von XPA und Rad52, die eine ähnliche Sequenz zum UNG Peptid aufweisen, mit RPA32C mittels NMR-Spektroskopie analysiert. Dies zeigte, dass RPA32C mit XPA und Rad52 ähnlich wie mit UNG interagieren kann und ermöglichten somit einen Einblick in die Protein-Protein-Interaktionen.^[113]

Eine weitere NMR Struktur konnte von der 18 kDa N-terminalen Domäne der RPA70 Untereinheit (Aminosäure 1-168) erhalten werden. Auch hier zeigte sich, dass die Reste 109-168 sehr unstrukturiert sind und vermutlich einen flexiblen Linker zwischen der DBD-F und DBD-A ausbilden. Die im N-Terminus enthaltene DBD-F (Aminosäure 8-108) besteht aus einem fünf-strängigen β -Fass. Des Weiteren besitzt das β -Fass an beiden Enden eine Art Deckel, der durch eine kurze Helix gebildet wird, und an einem Ende zwei Loops. Diese beiden Loops, die aus einem Lysin- und fünf Arginin-Resten bestehen, bilden eine basische Furche, um möglicherweise Interaktionspartner, die ein saures Motiv bzw. Oberfläche besitzen, zu binden (Abbildung 6b).^[114]

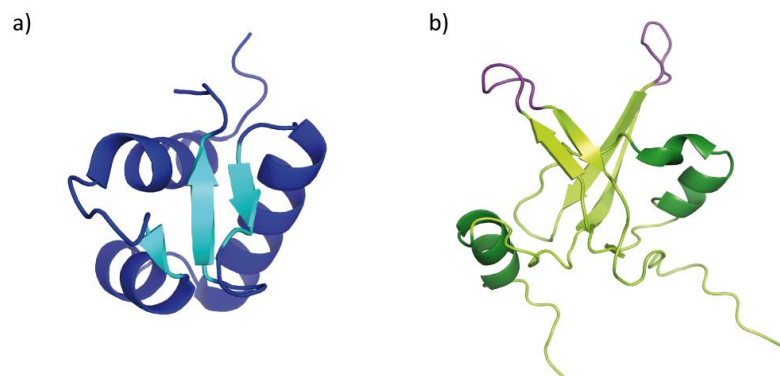


Abbildung 6: NMR-Bändermodelle von RPA32C und RPA70N. a) **RPA32C.** Die drei α -Helices vom humanen RPA32C sind in dunkelblau und die drei β -Stränge sind in hellblau hervorgehoben. (PDB: 1DPU). b) **RPA70N.** Das β -Fass vom humanen RPA70N ist in hellgrün, die zwei α -Helices sind in dunkelgrün und die zwei Loops sind in lila hervorgehoben. (PDB: 1EWI).

Weiterhin konnte eine NMR-Struktur des humanen Volllängen-Proteins RPA erhalten werden, welche Einblicke in der Dynamik und strukturelle Zusammensetzung ermöglichten. Dabei konnten über 350 von 550 erwarteten Signale der Domänen DBD-A/-B/-F, RPA32C und RPA32N im NMR identifiziert werden. Interessanterweise sind die Signale der DBD-A/-B und RPA70N identisch zu denen der einzelnen Domänen im NMR, was vermuten lässt, dass die Domänen ohne DNA strukturell unabhängig voneinander sind. Außerdem konnten keine Signale für den Trimerisierungskern (DBD-C/-D und RPA14) detektiert werden, obwohl dieser alleine im NMR gute Signale aufzeigt. Aufgrund dessen lässt sich vermuten, dass die zusätzlichen fünf Domänen die Rate des Taumelns (Bewegung) von dem Trimerisierungskern verlangsamen. Anschließend wurde das Volllängen RPA mit ssDNA analysiert und die erhaltenden Signale mit dem Spektrum des alleinigen Proteins verglichen. Dabei stellte sich heraus, dass die DBD-F und RPA32C keine Rolle in der Bindung an die ssDNA spielen und somit unter anderem frei für Protein-Protein-Interaktionen bleiben. Darüber hinaus konnte eine strukturelle Umlagerung der DBD-A/-B mit dem Trimerisierungskern detektiert werden. Folglich konnten *Brose et al.* die ersten Beobachtungen von RPA machen, das sich nach DNA-Bindung dynamisch umlagert.^[64]

1.3.5.3. Kristallstrukturen von RPA

Es konnten mehrere Kristallstrukturen der humanen RPA Untereinheiten, die zur Aufklärung der strukturellen Zusammensetzung von RPA beitrugen, durch Röntgen-Kristallstrukturanalyse erhalten werden. Eine davon ist die 2.5 Å RPA Kristallstruktur der DBD-A und der DBD-B von *Bochkareva et al.* (Abbildung 7a).^[73] Diese zeigt, dass die beiden Domänen aus je einem funktional unabhängigen OB-Motiv bestehen, eine N-terminale Verlängerung haben und die DBD-B eine zusätzliche C-terminale Helix besitzt. Die DBD-A und DBD-B sind in einer Tandem-Orientierung angeordnet und werden durch einen verlängerten Linker verbunden. Dieser Linker kann verschiedene Konformationen einnehmen und ist ohne DNA sehr flexibel.^[73]

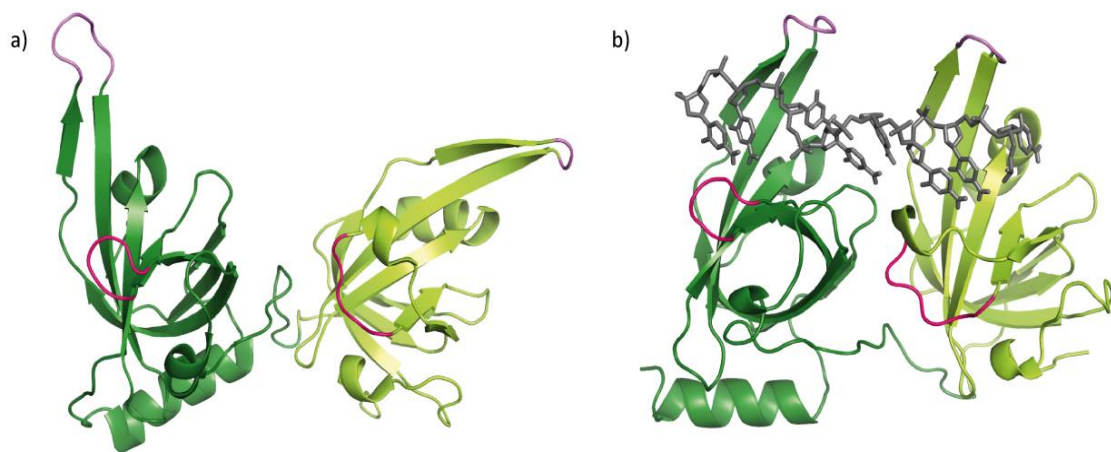


Abbildung 7: Vergleich beider Kristallstrukturen der humanen DBD-A und DBD-B von RPA70, ohne (a) und mit (b) DNA. Die DBD-A ist in hellgrün, die DBD-B ist in dunkelgrün und die DNA ist in grau hervorgehoben. Der Loop L12 ist in lila und der Loop L45 ist in pink markiert. (PDB: 1FGU, 1JMC).

Davor konnten *Bochkarev et al.* schon die DBD-A und DBD-B im Komplex mit ssDNA, einem Poly-dC₈-Oligonukleotid, kristallisieren (2.4 Å, Abbildung 7b).^[76] Bei Vergleich beider Strukturen wird ersichtlich, dass die Tandem-Orientierung durch die ssDNA Bindung stabilisiert wird, während die Struktur ohne DNA verschiedene Konformationen durch den flexiblen Linker einnehmen kann. Zudem ist durch Bindung der DBD-A und DBD-B an die ssDNA erkennbar, dass diese sich umorientieren, der Linker zwischen beiden Domänen stabilisiert wird und sich die Bindungsoberfläche für die DNA umlagert, um die DNA fest zu binden. Die DBD-A berührt das 5'-Ende und die DBD-B das 3'-Ende der DNA. Zudem tunnelt sich die DNA durch einen engen Kanal im oberen Bereich beider DBDs, wobei sie mehrere *stacking*- und hydrophobe Wechselwirkungen mit zwei Loops (L45 und L12) beider Domänen eingeht. Dieser DNA-Kanal entsteht durch die β_1 - und β_2 -Stränge, indem diese eine β -Haarnadel bilden, der sich um die DNA wickeln kann. Der L12 Loop weist basische Aminosäuren auf, die mit den sauren Phosphatgruppen der DNA interagieren können, und

die saure Bindungsfurche, die durch den Loop L45 gebildet wird, kann mit den DNA-Basen interagieren.^[73, 76] Des Weiteren kann die DBD-A mehr Wechselwirkungen mit der DNA eingehen als die DBD-B. Zusätzlich binden beide Domänen 3nt der DNA in der Kristallstruktur. Schlussendlich ist davon auszugehen, dass sich die Konformation der Domänen, besonders der Loops, durch die DNA-Bindung drastisch verändern.^[76, 110]

Neben diesen beiden Kristallstrukturen konnte erneut *Bochkarev et al.* die 2.5 Å Kristallstruktur von RPA32 (nur die DBD-D) zusammen mit RPA14 erhalten (Abbildung 8). Jene Struktur weist für beide Domänen ein OB-Motiv mit einer N-terminalen Verlängerung und einer C-terminalen Helix auf. Das OB-Motiv, besonders das von RPA32, ähnelt strukturell dem OB-Motiv von RPA70AB. Das Heterodimer interagiert durch eine Helix-Helix-Interaktion des C-Terminus miteinander.^[115]

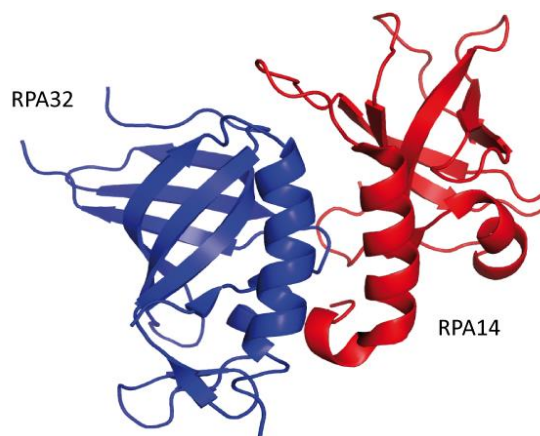


Abbildung 8: Kristallstrukturen des humanen RPA32/14 Dimers als Bändermodell. RPA32 ist in dunkelblau und RPA14 in rot dargestellt. (PDB: 1QUQ).

Deng et al. konnten 2007 die Kristallstruktur vom Volllängen RPA32 Protein mit RPA14 erhalten.^[116] Das OB-Motiv beider Domänen ist ähnlich zu der schon zuvor veröffentlichten Struktur. Dennoch war der N-Terminus sowie die wh-Domäne, die wahrscheinlich ungeordnet vorliegen, sehr schlecht aufgelöst und nehmen mehrere Positionen in der Kristallstruktur ein. Vermutlich kann die wh-Domäne mehrere Konformationen einnehmen, um die Protein-Protein-Interaktion mit verschiedenen Proteinen zu ermöglichen.^[116]

Die 2.8 Å Kristallstruktur des Trimerisierungskern von RPA (DBD-C/-D und -E) konnte von *Bochkareva et al.* im Jahr 2002 gelöst werden (Abbildung 9).^[58] Alle drei Domänen haben ein OB-Motiv mit einer C-terminalen α -Helix. Die Trimerisierung wird durch die C-terminalen α -Helices, die ein dreifaches-Helixbündel durch hydrophobe Interaktionen ausbilden, erreicht.^[58] Ansonsten kann die DBD-C, neben dem dreifach-Helixbündel, noch mit dem N-Terminus der DBD-D interagieren, wobei keine weiteren Interaktionen zu RPA14 ersichtlich sind. Beim Vergleich dieser Struktur mit der Kristallstruktur von RPA70AB bilden die Loops L12 und L45 der DBD-A und DBD-B eine tiefe ssDNA Bindungsfurche, die positiv geladen ist. Wohingegen die Loops von RPA32 und RPA14 wesentlich

kürzer sind und eine eher flache Bindungsfurche mit einer weniger positiv geladenen Oberfläche ausbilden. Die Bindungsfurche von DBD-E ist sogar negativ geladen und besitzt positiv geladene Loops. Dies könnte erklären, warum die DBD-D und DBD-E eine schlechtere Bindungsaffinität zur ssDNA haben als die anderen DBDs und wahrscheinlich daher die ssDNA mit ihrem OB-Motiven anders binden können. Die Bindungsfurche von DBD-C ist strukturell vergleichbar mit denen von RPA70AB und kann die ssDNA möglicherweise in einer ähnlichen Weise wie RPA70AB binden. Zudem besitzt die DBD-C ein Zink-Ribbon-Motiv, welches zwischen den β_1 - und β_2 -Strang eingebettet ist. Dieses Motiv wird durch drei antiparallelen β -Strängen gebildet, wobei das Zinkion durch vier Cysteine koordiniert wird. Zusätzlich besitzt die DBD-C ein Helix-Turn-Helix Motiv. *Bochkareva et al.* postulieren, dass dieses Motiv an der Bindung von dsDNA verantwortlich sein könnte.^[58, 116]

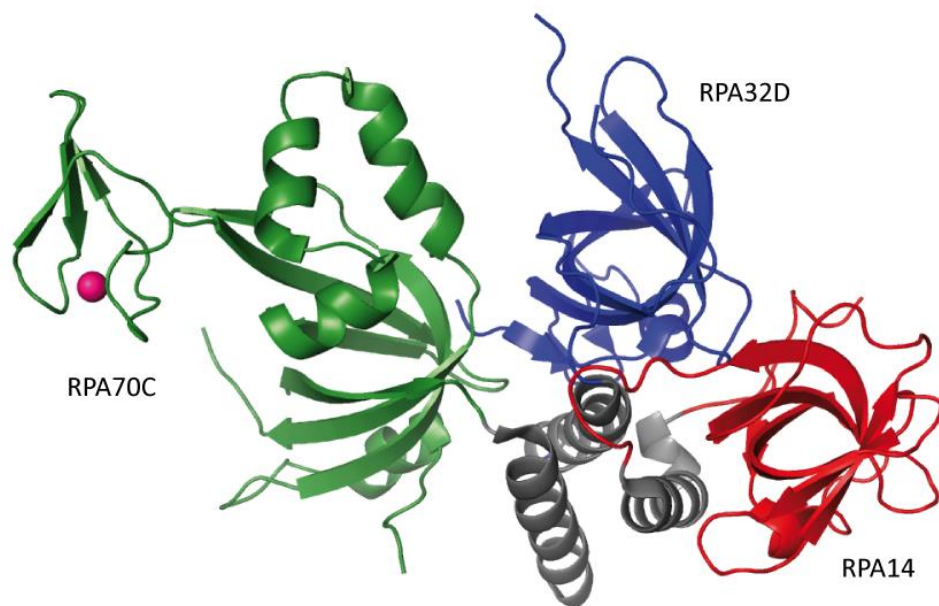


Abbildung 9: Kristallstruktur des humanen Trimerisierungskern von RPA als Bändermodell. Die DBD-C Domäne ist in dunkelgrün, die DBD-D Domäne ist in dunkelblau und RPA14 in rot dargestellt. Das dreifach-Helixbündel ist in grau und das Zinkion in pink markiert. (PDB: 1L1O).

Erst im Jahr 2012 konnten *Fan et al.* die 2.8 Å Kristallstruktur von RPA aus dem Pilz *Ustilago maydis* mit ssDNA (32nt) erhalten (Abbildung 10).^[117] Für diese wurde eine verkürzte Version von RPA verwendet, indem die N-terminale Domäne von RPA70 und die wh-Domäne von RPA32 entfernt wurden. Die Strukturen der einzelnen Domänen im Vergleich zu den humanen Kristallstrukturen sind sehr ähnlich. Es wird in der Kristallstruktur ersichtlich, dass der RPA-ssDNA Komplex eine quartäre Struktur einnimmt, in dem die ssDNA in einer U-Form von 25nt vorliegt. Die restlichen Nukleotide waren nicht gut genug aufgelöst. Dabei liegt RPA70 und RPA32 in einer Ebene, wohingegen RPA14 hinter dieser Ebene liegt. Der BC Linker interagiert mit DBD-B, DBD-C, ssDNA und bilden zusammen eine Vier-Wege-Schnittstelle, wodurch die Quartärstruktur stabilisiert wird. Zusätzlich wird die Quartärstruktur durch die Interaktion von DBD-A mit der C-terminalen Helix der DBD-C stabilisiert. Die DBD-A bindet 4nt am 5'-Ende der ssDNA und die DBD-B darauffolgend 5nt. Die nächsten 3

Nukleotide befinden sich genau am untersten Bereich der U-Form und interagieren mit dem BC-Linker, der DBD-B und der DBD-C Domäne. Die DBD-C bindet die nächsten 8nt und die letzten 5nt binden die DBD-D. Die DBD-A und DBD-B binden die DNA durch eine Furche in ihrem OB-Motiv, indem hauptsächlich vier aromatische hydrophobe Reste der Furche *van der Waals* Wechselwirkungen mit der ssDNA oder *stacking*-Interaktionen mit den Basen eingehen.^[117] Interessanterweise interkaliert der BC-Linker, der aus zwei β -Schleifen besteht, teilweise in die Furche der DBD-B, sodass die DBD-B eine andere Konformation als in dem 8nt Bindungsmodus (DBD-A, DBD-B mit 8nt ssDNA) von *Bochkarev et al.* einnimmt.^[76, 117] Zudem geht der BC-Linker viele Wasserstoffbrückenbindungen und *van der Waals* Wechselwirkungen mit der DBD-B, DBD-C und mit der ssDNA ein. Die DNA-Bindungsfurche der DBD-C ist wesentlich länger und breiter als die der DBD-A und DBD-B. Hinzu kann die Zinkdomäne der DBD-C mit dem Phosphodiester-Rückgrat interagieren. Dennoch hat die DBD-C eine geringere DNA-Affinität als DBD-A und DBD-B. *Fan et al.* vermuten, dass die DBD-C eine langsamere *on-rate* für ssDNA gegenüber der DBD-A und DBD-B besitzt. Dazu weist die DBD-D eine recht enge DNA-Bindungsfurche auf und geht nur wenige Wechselwirkungen mit der ssDNA ein, welches für die schwächere DNA-Bindungsaffinität dieser Domäne spricht. Bei einem Vergleich der Kristallstrukturen vom 8nt und 30nt Bindungsmodus der DBD-A und DBD-B unterscheiden sich diese in der Orientierung der Domänen und sowie in den Protein-DNA Kontakten.^[117]

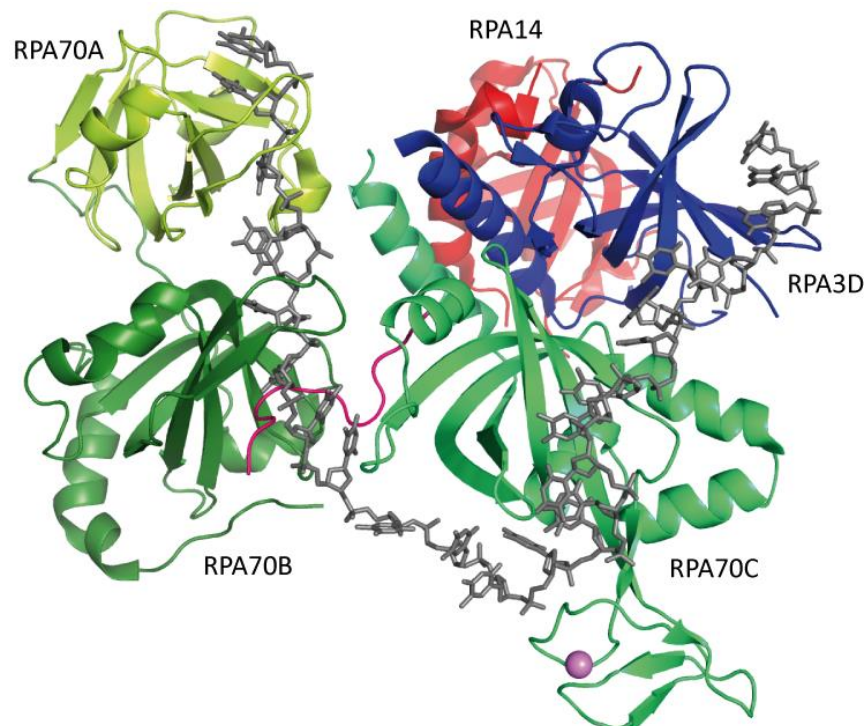


Abbildung 10: Darstellung der Kristallstruktur von RPA aus *Ustilago maydis*. Die DBD-A ist in hellgrün, die DBD-B ist in dunkelgrün, die DBD-C Domäne ist in mittelgrün, die DBD-D Domäne ist in dunkelblau und RPA14 in rot dargestellt. Der BC-Linker ist in pink, die ssDNA in grau und das Zinkion ist in lila markiert. (PDB: 1L1O).

Durch diese Kristallstruktur stellen *Fan et al.* die Hypothese auf, dass der BC-Linker eine wichtige strukturelle Rolle beim Übergang des 8nt Bindungsmodus zum 30nt Bindungsmodus spielt, indem dieser die DBD-C richtig positionieren kann, um die Vier-Wege-Schnittstelle aus DBD-B, BC-Linker, ssDNA und DBD-C zu bilden. Des Weiteren postulieren sie, dass die ssDNA Bindung von RPA oder die RPA-Protein-Interaktionen durch einen allosterischen Mechanismus erfolgt. Denn durch die Ausbildung der Vier-Wege-Schnittstelle wird die Zugänglichkeit der Protein-Protein-Interaktion Seiten der DBD-A und DBD-B verändert. Dieses könnte die Regulierung der DNA-Bindung bzw. DNA-Trennung von RPA sowie die Assemblierung und Disassemblierung anderer Proteine an RPA erklären.^[117] In Zukunft wäre die Struktur des humanen Volllängen-Proteins RPA mit DNA oder mit einem Interaktionspartner eine Bereicherung, um die Dynamik und Interaktion von RPA zu verstehen.

1.4. XPA und RPA

1.4.1. Interaktion von XPA und RPA

Die genaue Funktion von XPA und RPA in der NER ist noch nicht genau geklärt. Vermutlich können sie unter anderem den NER Komplex zusammen stabilisieren. Dennoch konnte bisher keine Struktur von XPA mit RPA erhalten werden, da beide Proteine viele strukturell ungeordnete Regionen besitzen. Trotzdem konnten die Interaktionsstellen beider Proteine entschlüsselt werden. In Abbildung 11 sind die Interaktionsdomänen veranschaulicht.^[118]

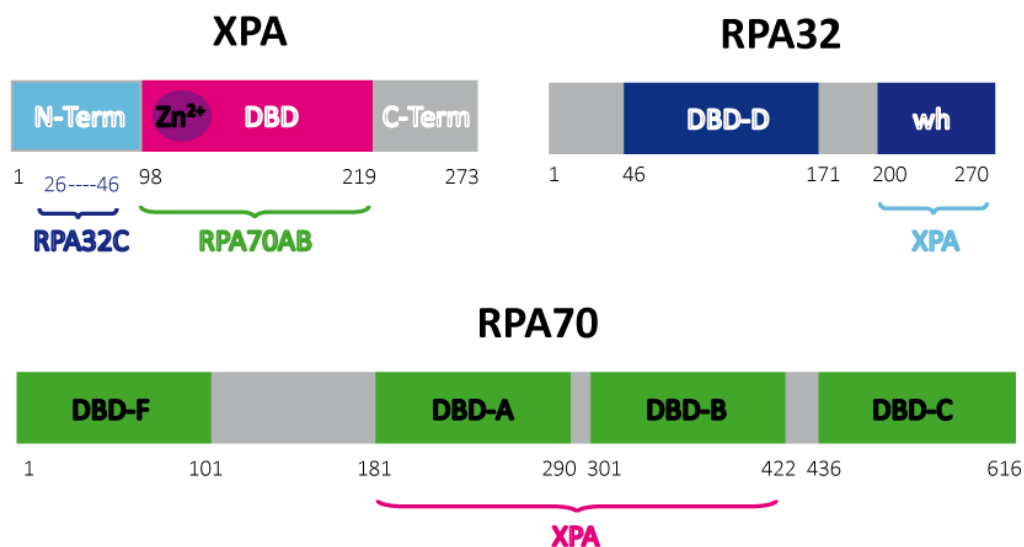


Abbildung 11: Schematische Darstellung der XPA und RPA Interaktionen. Der N-Terminus von XPA ist in hellblau, das Zinkfinger-Motiv in lila und die DBD in pink dargestellt. Die Domänen der RPA70 Untereinheit sind in grün und die RPA32 Domänen sind in blau markiert. Die mit der Klammer markierten Abschnitte interagieren mit den jeweils darunter befindlichen Protein.^[118]

XPA und RPA können durch zwei Interaktionspunkte miteinander interagieren. Dabei ist die Hauptinteraktion zwischen beiden Proteinen der N-Terminus von XPA (Aminosäure 29-46) und der C-Terminus von RPA32 (Aminosäure 204-270). Eine weitere schwächere Interaktion erfolgt durch die DBD von XPA (Aminosäure 98-219) mit den DBD-A/-B von RPA 70 (Aminosäure 181-422). Jedoch ist die genaue Interaktionsregion der DBD von XPA mit RPA70AB noch nicht genau geklärt.^[112, 118-120] Denn zum einen zeigten NMR-Titrationsstudien, dass das Zinkfinger Motiv von XPA involviert sein könnte. Im Gegensatz dazu zeigten *pull-down* Experimente und XPA-Mutanten, dass der C-Terminus der DBD von XPA eine Rolle bei der Interaktion mit RPA70AB spielt. Vor allem die Lysine K141 und K179 von XPA sind für die Wechselwirkung mit RPA70AB wichtig.^[98, 118, 121-123] Zudem ist es sehr interessant, dass die Bindungsregionen der XPA-DBD sowohl mit DNA als auch mit RPA überlappen, welches auf eine kooperative DNA Bindung zwischen RPA und XPA hinweisen könnte.^[121] Des Weiteren sind beide RPA Bindungsregionen von XPA durch einen großen, ungeordneten Linker verbunden.^[124]

1.4.2. Studien von XPA und RPA

Bis heute werden die genaue Funktion und das Zusammenspiel von XPA und RPA in der NER untersucht. Es ist bekannt, dass RPA den ungeschädigten DNA-Strang in 5'-3'-Richtung in der NER bindet. Daneben bindet XPA die DNA-Verzweigung in der NER an der 3'-Seite oder an der 5'-Seite, wobei die Bindung an der 3'-Seite eher angenommen wird.^[118, 125-126] Zudem konnte belegt werden, dass XPA bevorzugt *bulky* DNA-Schäden, vor allem zusammen mit RPA, bindet. Interessanterweise präferiert ein XPA/RPA-Dimer die Bindung an geknickte DNA.^[127-129] Auch zeigten *in-vitro* Studien, dass DNA-Schäden besser repariert werden, wenn XPA mit RPA vorinkubiert wird.^[130] Weiterhin kann XPA die Entwindung des DNA-Duplex durch RPA inhibieren, indem es wahrscheinlich beide Verzweigungen in der NER-Blase stabilisiert.^[131-132]

In den letzten Jahren wurde angenommen, dass XPA und RPA eine Rolle bei der Schadenserkennung bzw. bei der Schadensverifizierung spielen, da XPA verschiedene DNA-Schäden bindet und die DNA-Bindungsaktivität durch die Interaktion mit RPA gestärkt wird.^[112, 118, 129-131] Zudem könnte XPA zusammen mit RPA die DNA nach Schäden, vor allem nach DNA-Knicks und entwundenen Stellen, absuchen. Dadurch könnten DNA-Schäden detektiert und weitere NER-Proteine rekrutiert werden.^[133] Neue Studien zeigen, dass XPA und RPA erst nach der Ausbildung der NER-Blase rekrutiert werden, in der sich wahrscheinlich ein Komplex aus XPA und RPA bildet.^[118, 124, 134] Dabei könnte XPC oder auch TFIIH bei der Rekrutierung von XPA oder des XPA/RPA-Komplexes involviert

sein.^[135-136] Des Weiteren könnte auch ein XPA Homodimer (XPA₂) in der NER, vor allem in der Schadenserkennung, von Bedeutung sein. Dabei wäre es besonders interessant, wie sich aus XPA₂ ein XPA/RPA/DNA-Komplex ausbilden kann.^[124] Yang *et al.* konnten in einer Titrationsstudie mittels Gelfiltration zeigen, dass sich ein XPA₂-RPA-Komplex ausbildet.^[137] Außerdem ist die XPA/RPA-Interaktion in der Dissoziation von XPC in der NER involviert. Eine weitere Funktion des XPA/RPA-Komplexes ist die Rekrutierung und Positionierung der Exzisionsnukleasen.^[118, 138] E. Fadda stellte die Hypothese auf, dass XPA ein Multiproteinkomplex ist, welcher als ein flexibles Gerüst fungiert und eine Art „Perlen auf einer Schnur-Motiv“ besitzt, da die Protein-Protein-Interaktionsbereiche durch lange, ungeordnete Regionen getrennt sind. Dadurch könnten die „Perlen“ kooperativ fungieren.^[124, 139] Daneben stellte E. Fadda die Theorie, basierend auf die Überlappung der DNA-Bindungsregionen mit den XPA/RPA-Interaktionsregionen, auf, dass sich ein tertiärer XPA/RPA/DNA-Komplex bildet. Abbildung 12 zeigt das Modell von E. Fadda.^[124]

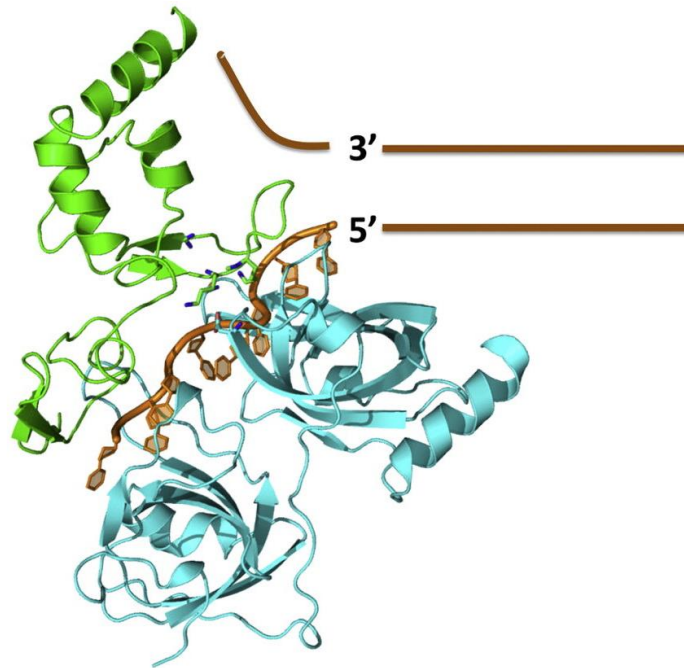


Abbildung 12: Ausschnitt des modellierten XPA₉₈₋₂₁₉-ssDNA-RPA70AB Komplexes an einer Y-förmigen DNA (DNA-Verzweigung) aus E. Fadda (DOI: <http://dx.doi.org/10.1016/j.csbj.2015.11.007>; Link zur Lizenz: <https://creativecommons.org/licenses/by/4.0/>).^[124] XPA₉₈₋₂₁₉ ist in grün, RPA70 ist in blau und die DNA ist in braun markiert.

Dabei soll ein Strang der DNA-Duplex sich durch beide Proteine tunneln und mit beiden interagieren. Um diese Theorie weiter zu verifizieren, hat E. Fadda den ternären Komplex mit dem Protein-Protein-Docking Programm *ClusPro* modelliert. Dafür diente die XPA₉₈₋₂₁₉ NMR-Struktur und die Kristallstruktur von RPA70AB mit DNA als Grundlage.^[76, 122] Das dabei entstandene Modell zeigt, dass die ungeschädigte ssDNA durch eine Art Tunnel, welcher durch beide Proteine ausgebildet wird, verläuft. Weiterhin beschreibt dieses Modell, dass die XPA₉₈₋₁₃₈ und die XPA₁₇₉₋₁₈₃ Regionen auch an

der Bindung von RPA70 und ssDNA beteiligt sind. Eine direkte Interaktion von RPA mit XPA kann durch die RPA Reste K179 und K167 erfolgen. Auch unterstützt dieses Modell, dass XPA an der 3'-Seite vom DNA-Schaden lokalisiert ist. Zudem würde die verlängerte DBD-Domäne von XPA beide Stränge der Y-förmigen DNA binden.^[124]

Die XPA/RPA/DNA-Interaktion mit beiden Vollängen Proteinen ist in Abb. 13 nach einem Modell von *E. Fadda* dargestellt.^[124] Dabei kann der unstrukturierte N-Terminus von XPA₂₉₋₄₆ mit einer globulären Domäne des C-Terminus von RPA32₂₀₄₋₂₇₀ an der 5'-Seite vom DNA-Schaden interagieren. Diese RPA32 Domäne bindet ssDNA nur in gewissem Maße. Daher könnte RPA auch als ein Ankerpunkt für XPA an der NER-Blase fungieren. Der unstrukturierte XPA Linker ist groß genug, um beide RPA-Bindungsmotive miteinander verbinden zu können.^[68, 124]

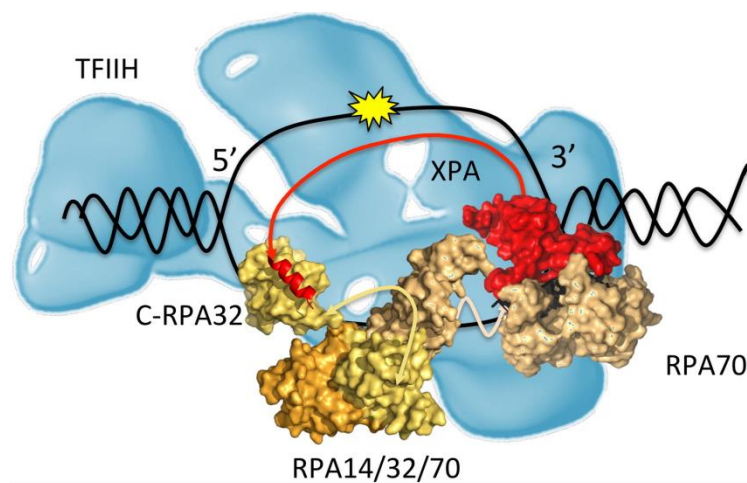


Abbildung 13: Ausschnitt eines modellierten XPA/RPA/TFIIH/DNA-Komplexes an einer NER-Blase aus *E. Fadda* (DOI: <http://dx.doi.org/10.1016/j.csbj.2015.11.007>; Link zur Lizenz: <https://creativecommons.org/licenses/by/4.0/>).^[124] XPA und der XPA Linker ist in rot, RPA70 ist in beige/gelb/orange, TFIIH ist in blau und die DNA ist in schwarz dargestellt.

Ein weiteres XPA/RPA-Modell, welches durch molekulares Docking mit dem Z-Dock Webserver und Strukturalignment erstellt wurde, wurde von *Hu et al.* eingeführt.^[140] Dafür wurde die XPA₉₈₋₂₁₉ NMR-Struktur und die Kristallstruktur von RPA70N verwendet.^[122, 141] *Hu et al.* konnten zeigen, dass das Zinkfinger-Motiv von XPA in eine V-förmige Tasche von RPA70N bindet. Dabei nimmt der Zinkfinger eine tetraedrische Struktur ein. Daneben sind vor allem die Reste Asp127-Asp134 von XPA an der Bindung zu RPA beteiligt. Außerdem führt der Austausch von Zink mit Nickel im Zinkfinger, wodurch dieser eine planare Konformation einnimmt, zu einer Schwächung der RPA-Bindung.

Letztendlich sind weitere Strukturanalysen, insbesondere *via* Kryo-Elektronenmikroskopie (Kryo-EM), des Multiprotein-Komplexes von hohem Interesse.

2. Aufgabenstellung

Die Nukleotidexzisionsreparatur ist unter anderem für die Entfernung von sterisch anspruchsvollen Basenaddukten verantwortlich. Der genaue Erkennungsmechanismus solcher Schäden ist bisher in der globalen Genomreparatur (GGR) unbekannt. Dabei besitzt das XPA Protein eine hohe Relevanz. Denn Defekte im XPA Protein führen zu den stärksten Phänotypen von *Xeroderma pigmentosum*. Zudem konnte noch nicht eindeutig geklärt werden, ob XPC oder XPA der initiale Schadenserkenner in der GGR ist.^[1]

Ziel dieser Arbeit ist es, mehr Kenntnis über den Erkennungsmechanismus und die Rolle von Rad14, das Hefehomolog des humanen Proteins XPA, in der NER zu erlangen. Die dabei ersten entstandenen Kristallstrukturen mit der DNA-bindenden Domäne von Rad14 und dG(C8)-AAF geschädigter DNA zeigten, dass die Bindung des Rad14 Proteins zur Ausbildung eines DNA-Knicks von ungefähr 70° führt, wobei sich der Schaden in der Knickstelle befindet. Zudem ist der Schaden in der DNA-Duplex interkaliert und die gegenüberliegende Base wird herausgeklappt, wodurch der Schaden als ein vollständiges Watson-Crick-Basenpaar fungiert. Die dabei möglichen π -stacking forces könnten für die Stabilisierung des DNA-Knicks eine bedeutende Rolle spielen. Dementsprechend ist die Untersuchung dieser π -stacking forces wichtig, denn diese könnten bedeutsam für die Ausbildung eines stabilen Erkennungskomplexes sein. Daher sollten verschiedene Schadensanaloga von AAF synthetisiert werden, in denen die aromatische Fläche verändert werden sollte. Nach Synthese von dG(C8)-Acetylphenyl-, dG(C8)-Acetylnaphthyl, dG(C8)-Acetylanthracenyl- und dG(C8)-Acetylpyrenyl-Phosphoramiditen, sollten diese durch DNA-Festphasensynthese in die jeweiligen Oligonukleotide eingebaut werden. Als nächstes sollten diese Oligonukleotide mit der DNA-bindenden Domäne von Rad14 kristallisiert werden, um weitere Aussagen über den Erkennungsmechanismus von Rad14 machen zu können. Zusätzlich sollte auch dG(C8)-AF-geschädigte DNA und verschiedene N^2 -geschädigte DNA Addukte mit Rad14 kristallisiert werden. Die Kristallstrukturen der N^2 -Addukte wären sehr interessant, da diese Art von Schäden von der TCR repariert werden.

Ein weiteres Ziel dieser Arbeit ist es, ein XPA/RPA/DNA-Komplex zu bilden und zu isolieren. Anschließend sollte durch Strukturanalysen wie Kristallisation oder Kryo-Elektronenmikroskopie die Struktur dieses Komplexes gelöst werden, um weitere Einblicke in den NER-Mechanismus zu bekommen. Zudem sollten verschiedene DNA-Substrate für die Komplexbildung überprüft werden, um das am besten geeignete Substrat zu finden. Nach erfolgter Expression und Aufreinigung von XPA und RPA sollten auch EMSA (*electrophoretic mobility shift assay*) Studien mit verschiedenen DNA-Substraten durchgeführt werden, damit Aussagen über die Bindungsaffinitäten getroffen werden können. Des Weiteren sollten *Crosslinking*-Studien der einzelnen Proteine beziehungsweise

des Komplexes durchgeführt werden, um die strukturelle Interaktion der Proteine zueinander zu analysieren.

3. Das Hefeprotein Rad14

3.1. Struktureller Einblick in die Erkennung von Cisplatin- und AAF-dG Schäden durch Rad14 (XPA)

S. C. Koch*, J. Kuper*, K. L. Gasteiger*, **N. Simon**, R. Strasser, D. Eisen, S. Geiger, S. Schneider, C. Kisker, T. Carell, *PNAS* **2015**, *112*, 8272–8277. „*Structural insights into the recognition of cisplatin and AAF-dG lesion by Rad14 (XPA)*“. (* geteilte Erstautorenschaft)

3.1.1. Einführung

XPA besitzt eine hohe Relevanz in der NER, denn Defekte in diesem Protein führen zu den stärksten Phänotypen von *Xeroderma pigmentosum*. Bisher wurde angenommen, dass XPA eine Rolle in der Schadenserkennung und –verifizierung sowie als Gerüstprotein spielen könnte. Dennoch ist die genaue Funktion von XPA bis heute noch nicht bekannt. Dieses Manuskript beschreibt die erste Kristallstruktur von Rad14, das Hefehomolog von XPA, mit C8-AAF-dG und Cisplatin-geschädigter DNA und zeigt dabei erste Einblicke in die Funktion von Rad14/XPA. Dabei wurde aufgedeckt, dass Rad14 die DNA-Schäden durch einen *sequence-independent shape-readout* Mechanismus erkennt, indem es die DNA auf Biegsamkeit und Flexibilität überprüft.

3.1.2. Autorenbeitrag

Die Hauptarbeit bei diesem Projekt lag bei *Sandra C. Koch*. Für dieses Manuskript führte ich die Bindungsstudien mit FITC-geschädigter DNA (15mer und 48mer) und Rad14/XPA in Zusammenarbeit mit *Sandra C. Koch* durch. Dafür reinigte ich auch die beiden Proteine Rad14 und XPA auf. Weiterhin habe ich die *Crosslinking*-Studien mit Rad14 und dem BIS(NHS)PEG₅ *Crosslinker* in enger Zusammenarbeit mit *Sandra C. Koch* durchgeführt. Dafür synthetisierte ich die 15mer und 37mer DNA mit einem C8-AAF-dG Schaden. Ebenfalls hierfür reinigte ich das Rad14 Protein auf.

Structural insights into the recognition of cisplatin and AAF-dG lesion by Rad14 (XPA)

Sandra C. Koch^{a,1}, Jochen Kuper^{b,1}, Karola L. Gasteiger^{a,1}, Nina Simon^a, Ralf Strasser^a, David Eisen^a, Simon Geiger^a, Sabine Schneider^c, Caroline Kisker^{b,2}, and Thomas Carell^{a,2}

^aCenter for Integrated Protein Science at the Department of Chemistry, Ludwig-Maximilians Universität München, 81377 Munich, Germany; ^bRudolf Virchow Center for Experimental Biomedicine, Institute for Structural Biology, University of Würzburg, 97080 Würzburg, Germany; and ^cDepartment of Chemistry, Technische Universität München, 85747 Garching, Germany

Edited by Suse Broyde, New York University, New York, NY, and accepted by the Editorial Board May 25, 2015 (received for review May 5, 2015)

Nucleotide excision repair (NER) is responsible for the removal of a large variety of structurally diverse DNA lesions. Mutations of the involved proteins cause the xeroderma pigmentosum (XP) cancer predisposition syndrome. Although the general mechanism of the NER process is well studied, the function of the XPA protein, which is of central importance for successful NER, has remained enigmatic. It is known, that XPA binds kinked DNA structures and that it interacts also with DNA duplexes containing certain lesions, but the mechanism of interactions is unknown. Here we present two crystal structures of the DNA binding domain (DBD) of the yeast XPA homolog Rad14 bound to DNA with either a cisplatin lesion (1,2-GG) or an acetylaminofluorene adduct (AAF-dG). In the structures, we see that two Rad14 molecules bind to the duplex, which induces DNA melting of the duplex remote from the lesion. Each monomer interrogates the duplex with a β -hairpin, which creates a 13mer duplex recognition motif additionally characterized by a sharp 70° DNA kink at the position of the lesion. Although the 1,2-GG lesion stabilizes the kink due to the covalent fixation of the crosslinked dG bases at a 90° angle, the AAF-dG fully intercalates into the duplex to stabilize the kinked structure.

XPA | Rad14 | NER | AAF | cisplatin

Defects of the nucleotide excision repair (NER) system cause the human disease xeroderma pigmentosum, which is characterized by hypersensitivity to sunlight, resulting from the inability of patients to repair UV-induced DNA lesions (1, 2). Eight XP complementation groups are known, of which seven are caused by mutations in genes involved in the NER process (3, 4). NER recognizes a large array of diverse lesions (5–9). It repairs UV- and cisplatin-induced intrastrand crosslinks (10, 11). As such, NER protects higher organisms from the harmful effects of sunlight but also establishes resistance against cisplatin therapeutics, which is a major problem associated with this type of chemotherapy (12). The NER system also repairs a wide variety of single base bulky DNA adducts formed by environmental carcinogens. NER is thus the most versatile known DNA repair system (5, 8, 9).

The structurally vastly different lesions are recognized either within a globally operating NER (global genome repair, GG-NER) (7, 13) or as part of a transcription coupled NER process (TC-NER), where the stalled RNA polymerase is the initial NER inducing signal (14–16). Several lines of evidence lead to the currently accepted hypothesis that the lesions are initially recognized by XPC with the help of UV-DDB (UV DNA damage binding protein) for cyclobutane pyrimidine dimer (CPD) lesions (17–20) followed by damage verification by XPD/TFIIH (21–24), XPA, and XPG (3) to finally assemble to the preincision complex (3, 25–27). This step seems to involve recruitment of the XPA protein, which is one of the proteins essential for both GG-NER and TC-NER (28). As such, mutations of the XPA proteins provide one of the strongest NER phenotypes (29). Although the function of most XP-proteins within the NER process is understood, the precise role of XPA has remained unclear. XPA is known to interact with other NER proteins such as RPA, TFIIH, and

ERCC1 (30–33) suggesting XPA to be a NER scaffold protein (3). Furthermore, detailed biochemical experiments showed that XPA binds to kinked DNA structures (34) and to DNA duplexes containing certain types of DNA lesions, such as cisplatin (10, 35, 36) adducts and bulky adducts (5, 37, 38). XPA was found to form a homodimer (38), and it was demonstrated that it forms a 2:1 complex with the respective DNA (37). Initial structural insights into the protein architecture were derived from an NMR structure of the DNA binding domain of XPA (39, 40). However, the precise structural basis of the interactions of XPA with its DNA substrates is not known (3, 28).

Here we present two structures of the *Saccharomyces cerevisiae* homolog of XPA, Rad14, in complex with DNA containing a cisplatin 1,2-GG intrastrand crosslink and a *N*-(deoxyguanosin-8-yl)-2-acetylaminofluorene (AAF-dG) bulky adduct, which are representatives of typical NER substrates (6). The obtained structures shine light on how one of the major NER proteins interacts with DNA.

Overall Structure of Rad14 Bound to 1,2-GG Cisplatin and AAF-dG Containing DNA

We first investigated binding of Rad14 to double-stranded DNA containing different lesions in a central position (ODN1-5, see Table S1) using band shift assays (Fig. S1). We confirmed complex formation with DNA containing a cisplatin 1,2-GG (10, 36, 41) and an AAF-dG lesion (37, 38) (Fig. S1 *D* and *E*). We next studied

Significance

Nucleotide excision repair (NER) is a versatile repair machinery able to protect organisms from DNA damage. Defective NER leads to diseases like xeroderma pigmentosum (XP). XPA is a central NER protein that interacts with DNA in an unknown fashion. Here we present two crystal structures of the yeast homolog of XPA, Rad14, in complex with two NER substrate lesions. Rad14 binds to the damaged DNA from both sides of the lesion. Binding creates a sharp kink of the duplex by 70°. Each protein inserts a hairpin loop into the duplex to induce partial melting around the lesion. The structures provide insight into the mechanism of how XPA binds to kinked and lesion-containing DNA.

Author contributions: C.K. and T.C. designed research; S.C.K., J.K., K.L.G., N.S., R.S., D.E., S.G., and S.S. performed research; S.C.K., J.K., K.L.G., R.S., D.E., S.S., C.K., and T.C. analyzed data; and S.C.K. and T.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. S.B. is a guest editor invited by the Editorial Board.

Data deposition: Atomic coordinates have been deposited in the Protein Data Bank, www.ebi.ac.uk/pdbe (PDB ID codes 5a3d for AAF-dG and 5a39 for cisplatin).

¹S.C.K., J.K., and K.L.G. contributed equally to this work.

²To whom correspondence may be addressed. Email: thomas.carell@cup.uni-muenchen.de or caroline.kisker@virchow.uni-wuerzburg.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1508509112/-DCSupplemental.

the binding stoichiometry with AAF-dG containing DNA (Fig. S1E) and confirmed the expected formation of a 2:1 complex with Rad14 (37, 38).

For crystallization studies, we used two truncated versions of Rad14 (Rad14_{188–302/306}), which represent the DNA binding domain. We confirmed that the truncated version of Rad14_{188–302} (Rad14t) retains the binding characteristics to DNA (Fig. S1G). Cocystals were obtained with Rad14t in complex with a 15mer DNA duplex containing a single AAF-dG in the central position and with a 16mer duplex with a central cisplatin 1,2-GG lesion (Fig. 1 C and D). We first determined the structure of selenomethionine containing Rad14_{188–306} featuring a C-terminal strep-tag in complex with the AAF-dG containing DNA, which provided

good structural data for the protein at 3.1-Å resolution. However, large parts of the DNA, including the area in which the damage was located, remained uninterpretable. We therefore substituted three thymidines with 5-iodo-uracils and replaced the tag. This new construct resulted in a different crystal form, of which the Rad14t moiety was solved by molecular replacement using the coordinates of the previously experimentally determined selenomethionine-labeled protein and the DNA was located by Fourier analysis. The structure was refined to a final resolution of 1.8 Å. In addition, the structure of the DNA-binding Rad14t fragment in complex with the 16mer DNA duplex containing a central cisplatin 1,2-GG lesion was determined to 2.8-Å resolution (Table S2).

The Rad14 structures show that the protein assumes an overall α/β -fold (Fig. 1B) quite similar to the NMR structure (Fig. S2A) of the human XPA homolog, indicating high structural conservation (40, 42). In agreement with the 2:1 stoichiometry determined by gel shift experiments (Fig. 3 F and G and Fig. S1E), we see in both structures that two Rad14 molecules bind to the DNA duplex above and below the respective lesion. This binding is shown for the 1,2-GG cisplatin containing DNA in Fig. 1C and for the AAF-dG in Fig. 1D. Interestingly, the Zn-finger domains of both proteins are not in contact with the duplex DNA.

In the structure with the 1,2-GG lesion (Fig. 1C), two Rad14 molecules bind to the undamaged segments of the DNA above and below the lesion. Each protein interrogates the duplex with a β -hairpin (residues 253–267), leading to stacking of Phe262 and His258 onto the last intact base pair on either DNA end and unpairing of the following bases (Fig. 1A). This intercalation on both ends creates a central 13-bp duplex recognition element with the lesion in the middle. The entire structure exhibits a pseudo C2 symmetry because the DNA is likewise bound in two different orientations, which differ by a 180° rotation around the central positioned lesion. Because the 1,2-GG lesion is a cross-link between two central dG bases, the lesion can be observed for each of the two possible DNA orientations in two different positions (in total four positions, see Fig. S3A). For clarity only one of the structures is shown in Fig. 1C. This observation already suggests that Rad14 binds lesion independently, mostly triggered by conformational effects. We used the anomalous signal of the Pt-atom to estimate the occupancies (DNA orientation 1: 21% and 27%; DNA orientation 2: 33% and 19%, see Fig. S3A). In the structure with the AAF-dG lesion (Fig. 1D), we observed that the DNA is also bound in two orientations with the AAF-dG lesion located in two different positions linked by the pseudo C2-symmetry axis. The two duplex orientations are occupied with an approximate 45–55% distribution as inferred from the anomalous signal of the iodine (Fig. S3B). Importantly, the 13mer DNA lesion recognition motif in both structures (AAF-dG and 1,2-GG) is fully base paired and the lesions are both not in a flipped out state. Instead, the structures exhibit a sharp kink of 70° at the position of the lesion showing that Rad14 does not recognize the lesion itself but likely senses the flexibility of the DNA created in both cases by the lesions (Fig. 2).

The DNA ends below and above the kinked 13mer recognition motif are unwound, they lack any H-bonding interactions and are partially disordered in the structure. The corresponding 5' ends of the DNA duplex at both ends of the structure are bound in a cleft created by the β -hairpin, the end of helix α_4 , and the anti-parallel β -sheet domain (Figs. 1B and 2A). The β -hairpin structure and this cleft are responsible for melting of the duplex above and below the 13mer recognition sequence. A recent DNA binding study by Sugitani et al. (43) suggests that the DNA binding properties of XPA are improved if the binding domain is extended at the C terminus by 20 amino acids (XPA_{98–239}), of which 9 provide basic side chains. The improved binding can be explained with the presented structure. Addition of the amino acids extends helix α_7 (Fig. 1D), which leads to improved packing against the DNA duplex

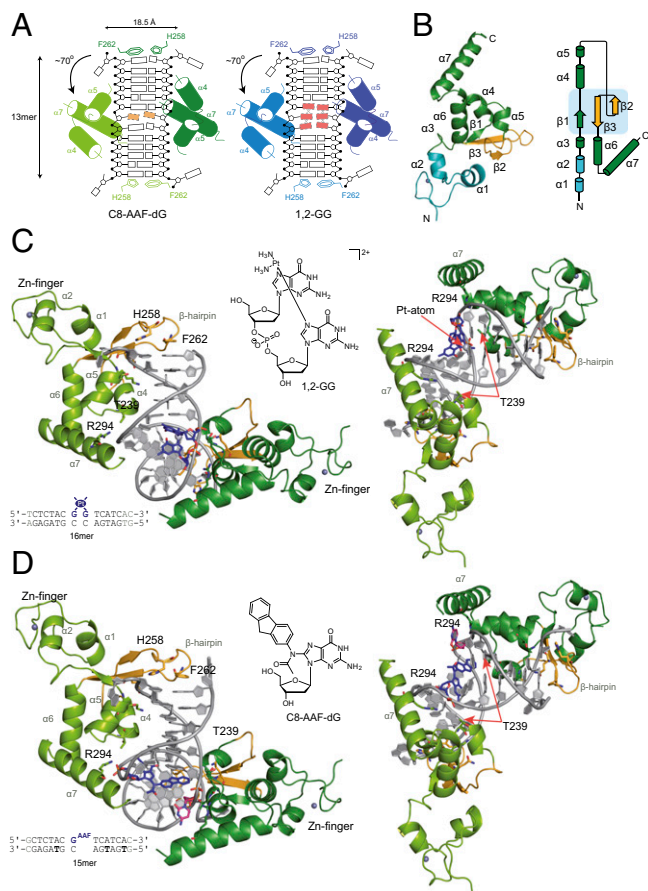


Fig. 1. Overall structures of the Rad14_{188–302}-DNA complexes. (A) Schematic diagrams of the Rad14-AAF-dG and Rad14-1,2-GG complexes showing the different positions where the lesions were observed in orange and red, with the C1'-C1' distance of the last base pair in the molten part of the structure. For the AAF-dG lesion (Left) two positions of the lesion and for the 1,2-GG lesion (Right) 4 positions of the lesion were observed (Fig. S3). For clarity, only one DNA orientation with the cisplatin lesion is shown here. (B) Schematic representation of the α/β -folding topology of Rad14_{188–302}, with the central 3-stranded β -sheet. (C) Ribbon diagrams of the Rad14-cisplatin lesion DNA complex with the structure of the lesion and the DNA sequence. Rad14 is shown in green and gold (β -hairpin), the DNA backbone is shown in gray, and the cisplatin 1,2-GG lesion is shown in blue. Residues important for DNA kinking (Thr239, His258, Phe262, and Arg294) are shown as sticks. (D) Ribbon diagrams of the Rad14-AAF-dG lesion DNA complex with the structure of the lesion and the DNA sequence. Rad14 is shown in green and gold (β -hairpin), the DNA backbone is shown in gray, and the AAF-dG lesion is shown in blue. Residues important for DNA kinking (Thr239, His258, Phe262 and Arg294) are shown as sticks. In the DNA sequences, the unpaired, partially disordered DNA bases are depicted in light gray. Thymidines replaced by 5-iodo-uracils in the AAF-dG complex are shown in bold.

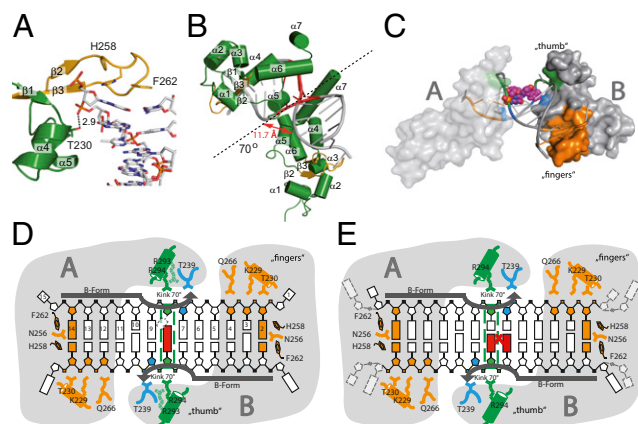


Fig. 2. Rad14t binds to lesion containing DNA by bending of the DNA. (A) Close up view of the β -hairpin structure with Phe262 and His258 stacking on top of the last base pair. The β -hairpin is shown in orange, Rad14t in green and the DNA as stick model. (B) Bending of the DNA duplex at the lesion position (red) by 70° reduces the P–P distance in the major groove to 11.7 Å. The dotted line represents the pseudo C2 symmetry axis. (C) Schematic representation of the bending process. The β -hairpin is shown in orange, the packing of $\alpha 7$ with Arg294 against the backbone is shown in green, and the stabilization of the bent by Thr239 is shown in blue. (D) Schematic representation of the interaction in the Rad14t-AAF-dG DNA (15mer duplex) complex showing the “fingers” domain mainly formed by the β -hairpin as well as residues Gln266, Lys229, and Thr230 and the “thumb” created by $\alpha 7$ with Arg294. Arg293 (light green) interacts with the backbone only in the AAF-dG structure. The dotted square indicates the dC base opposite the AAF-dG lesion, for which electron density is missing likely due to flexibility. (E) Schematic drawing of the interactions between Rad14t and cisplatin 1,2-GG DNA (16mer duplex) using the same color code as in D.

backbone with additional charge interactions between the negatively charged backbone and positive amino acids.

DNA Bending and the Lesion Recognition Motif

The melting process, which is induced by the β -hairpin irrespective of the lesion (Figs. 14 and 24), separates the bases at the duplex ends by about 18.5 Å (C1'–C1' distance base pair 1 and 15). The most prominent feature of the structures is the formation of a sharp 70° kink at the central position (Fig. 2 B and C) where the lesions are situated. Bending occurs into the major groove and is characterized by a reduction of the interstrand phosphodiester distance at the closest point of the concave side to 11.7 Å compared with 16.8 Å in B-DNA. The protein–DNA interfaces (Fig. 2 D and E) that enable these processes are composed of residues from the β -hairpin (Asn256, His258, Phe262, and Gln266), $\alpha 4$ (Lys229 and Thr230), the loop between $\alpha 4$ and $\alpha 5$ (Thr239), and $\alpha 7$ (Arg293 and Arg294). The most prominent feature of the aforementioned interfaces is the β -hairpin (Fig. 24), which establishes the majority of protein–DNA interactions, representing the anchor point for the bending process. The DNA is held in place by intercalation of Phe262 and His258 and further interactions established by Lys229, Thr230, and Gln266, which hydrogen bond to the phosphate moieties of base 2 and 3 in the 5' strand. His258 forms a hydrogen bond with the phosphate moiety of the base, which gives rise to partially distorted stacking interactions. Asn256 and Gln266 contribute hydrogen bonds to base 2 and the phosphodiester of base 4 (Fig. 2 D and E). These and all other interactions (with the exception of Asn256) observed in the complete protein DNA interface are exclusively achieved through phosphate backbone interactions, which ensure a sequence independent recognition mechanism. They can be described as interrogating “fingers” that tightly bind the duplex starting 4–5 base pairs away from the lesion. A second “thumb” like interaction with the DNA backbone is established by helix $\alpha 7$, which packs against

the backbone close to the lesion. As observed by Sugitani et al., extending this helix would likely result in even tighter binding (43). The “thumb” holds the DNA with the help of Arg294, forming a charge interaction with the phosphodiester. The “thumb” is further supported by an interaction established by Thr239 located between $\alpha 4$ and $\alpha 5$. In the AAF-dG structures (see below), we observe in addition that Arg293 keeps the backbone in place. The combined action of the intercalated Phe262 and His258 residues and the arginines in the $\alpha 7$ –“thumb” forces the DNA to bend into the major groove. Because the bent conformation is established by only two interactions, with Arg294 and Thr239, the structure supports the hypothesis that Rad14 binds to DNA structures that are easily bendable or which already adopt a bent conformation before Rad14 binding. In this respect, our structures also explain the observed binding of XPA/Rad14 to DNA with bulges (34) because such DNA structures are also kinked (44). The structures furthermore explain the observed cooperativity of the Rad14/XPA binding event that is also visible in our gel shift experiments (Fig. 3 F and G and Fig. S1E) (37). In our model, the first Rad14 might bend the duplex slightly, thus facilitating the second Rad14 binding step, which provides the ability to form the sharp kink.

Analysis of the (Rad14)₂-DNA Complex in Solution

To prove that the crystallographically observed structures also exist in solution and to investigate the possibility that complex formation involves binding of Rad14 just to the DNA ends, we performed protein–protein crosslinking experiments. Analysis of the crystal structures shows that Lys233 of each Rad14 points toward each other at a distance of 21.5 Å (Fig. 3A). We therefore prepared a longer DNA 19mer duplex (ODN8, see Table S1) with a central AAF-dG lesion and incubated it with Rad14. To this solution we added the reactive Bis(NHS)PEG₅ crosslinker (Fig. 3B), which places the reactive ester groups at a distance of ~21.7 Å assuming an extended conformation. Analysis of the reaction by gel electrophoresis (Fig. 3B) confirmed the formation of a defined (Rad14)₂-DNA species. Interestingly, formation of only one crosslinked species even in a concentration dependent fashion was observed arguing that the distance between the two Rad14 molecules is around 21.5 Å despite the increased length of the duplex. To prove that the crosslinker has bridged the Lys233 residues of the two monomers we next digested the crosslinked (Rad14)₂ with trypsin and analyzed the fragments by HPLC-MS/MS. The data show a parent ion signal with the exact molecular weight ($m = 2115.02$ Da) of the expected crosslinked dipeptide species (Fig. 3C) that was further characterized as TECKEDY (marked in red in Fig. 3A) by MS sequencing (Fig. 3D and E). We also detected the typical fragmentation pattern of the ethylene glycol (PEG) crosslinker with $\Delta m = 44.03$ Da confirming the presence of the crosslinker.

Despite the clear results, it cannot be fully excluded that the crosslinker reacts with endbound Rad14 with the DNA being in an unusual bent conformation. To exclude this possibility, we performed a crosslinking experiment side by side with AAF-dG containing duplexes of different lengths (15mer ODN6 and 37mer ODN9, Fig. 3H). Despite the length difference, similar crosslinking efficiencies were observed, which argues against this possibility. The crystallographically determined binding mode is furthermore supported by the observation that Rad14 also binds to longer duplexes only in the presence of a lesion (Fig. 3G). Undamaged duplexes are not recognized.

To further investigate the importance of the observed β -hairpin intercalation, we performed fluorescence depolarization studies with Rad14t and a Rad14t variant in which we replaced one of the intercalating residues of the β -hairpin, Phe262, by an alanine (Rad14t_{F262A}). For this study FITC-dU containing duplexes of different lengths (15mer ODN11 and 30mer ODN10, see Table S1) were used, in which the fluorescein, needed for the detection, functions as the lesion. It is well known that Rad14 binds the

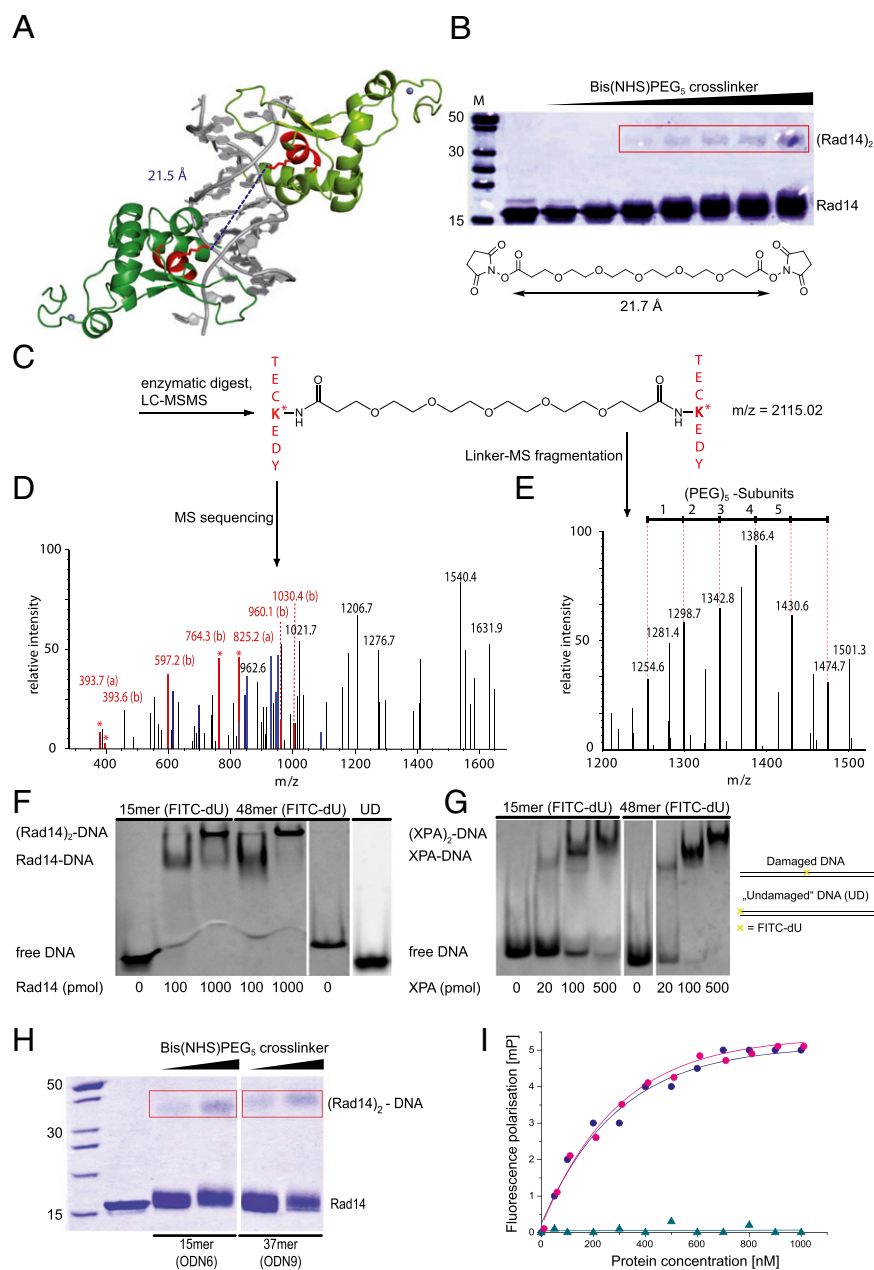


Fig. 3. Crosslinking experiment of the Rad14t-DNA complex and mass spectrometric analysis of the digested protein dimer. (A) Two Rad14t proteins (light and dark green) bind to one DNA strand containing the AAF-dG lesion (gray). The two lysines at a distance of 21.5 Å in the peptide sequence TECKEDY are highlighted in red. (B) SDS gel of the protein-DNA mixture incubated with increasing amounts of the Bis(NHS) PEG₅ crosslinker. (Rad14)₂ bands are boxed in red. The bands were cut out and subjected to enzymatic digestion. (C) Crosslinked peptide sequences after enzymatic digest of the proteins. (D) The MS/MS-spectrum created in an attached HCD-cell of the mass spectrometer reveals the peptide sequence enabling peptide identification (the a- and b-ion series is shown in red, x- and y-ion series is shown in blue). (E) Mass spectrometric analysis revealing the PEG fragments from the crosslinker. The PEG chain fragments with a typical $\Delta m = 44$ Da. Electrophoretic mobility shift assay (EMSA) of Rad14t (F) and XPA (G) proteins with a central FITC-dU lesion containing and undamaged DNA (UD; endstanding FITC labeled DNA ODN3 and ODN5, see Table S1). (H) SDS PAGE of the protein-DNA mixture incubated with increasing amounts of the Bis(NHS)PEG₅ crosslinker. A 15mer (ODN6) and a 37mer (ODN9) DNA duplex containing an AAF-dG lesion were used. (Rad14)₂ bands are boxed in red. (I) Fluorescence depolarization data showing the DNA binding properties of Rad14t [blue dots: 30mer (ODN10) and purple dots: 15mer (ODN11)] and Rad14t_{F262A} (cyan triangles) to a central FITC-dU lesion containing DNA duplexes.

fluorescein-containing base with significant affinity (see below) (45). Formation of the lesion recognition complex with FITC-dU and the full length proteins Rad14t and XPA was demonstrated by a gel shift experiment (Fig. 3 F and G). As depicted in Fig. 3 I, the F262A point mutation fully abolished the lesion recognition ability of Rad14t, thus solidifying the model that the intercalation of the β -hairpin into the duplex is crucial for lesion recognition.

Discussion

In contrast to other DNA repair mechanisms, NER successfully recognizes structurally vastly differing lesions. Although DNA binding of the NER proteins XPC (46–48) and XPE (49, 50) are structurally understood, for XPA, which exhibits the strongest NER phenotype, it is only known that it interacts with kinked DNA structures such as DNA containing bulges and also in special cases with damaged DNA (10, 51, 52). Structural information about the binding process is lacking. Our structure of the XPA homolog Rad14 in complex with lesion containing DNA now

uncovers the mechanism of how XPA interacts with the DNA. The obtained crystallographic results are fully consistent with previous biochemical data that showed binding of XPA to kinked DNA structures (34). Importantly, Rad14/XPA, like Rad4/XPC (18), does not bind the lesion directly but recognizes weakened DNA duplex structures. We observe that Rad14/XPA binding goes in hand with formation of a sharp kink at the lesion site by 70°. DNA binding involves two Rad14 proteins. Each protein inserts a β -hairpin exactly 6 base pairs away from the lesion, which generates a 13mer recognition motif. The lesion is not in a flipped out state but stays inside the 13mer duplex motif to stabilize the sharply bent structure. Bending occurs into the major groove, which in the case of the cisplatin 1,2-GG lesion means toward the Pt-atom (Fig. 44). The intercalation of the β -hairpin “fingers” separates the DNA strands to form single stranded regions above and below the 13mer Rad14 binding motif to which other NER proteins may bind. This DNA binding mode could be one element that allows XPA to interact with many other NER factors (31–33).

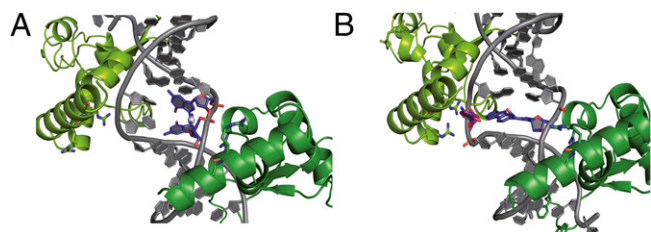


Fig. 4. Rad14t recognizes lesions that can stabilize kinked structures. (A) Close-up view of the 1,2-GG lesion with the cisplatin unit (blue) intercalated between adjacent base pairs (gray) to stabilize the kinked structure formed by (Rad14t), (green). Arg294 is depicted in stick representation with the nitrogens in blue. (B) Close up view of the AAF-dG lesion in complex with Rad14t using the same color code as in A. The flipped out dC was modeled into the structure and is shown in red/blue.

The relatively few interactions between Rad14 and the duplex suggests that Rad14 binds best to already prekinked DNA structures (34).

This explains why Rad14 is able to bind to DNA containing bubbles, bulges and Y-junctions (34). All these substrates can form easily kinked DNA structures (52, 53). CPD lesions, on the other hand, are not recognized by Rad14, likely, because these lesions are less efficiently accommodated in a sharply kinked structure. Indeed, it is known that CPD lesions hardly influence the duplex structure and its stability (44). More importantly, however, is the fact that the stiff four-membered ring cyclobutane structure of the CPD lesion points into the major groove, which might block bending into this groove.

For single-base bulky adduct lesions such as AAF-dG, it is unclear to which extent they prekink the duplex. For AAF-dG and AF-dG lesions it was suggested that they exist in two different orientations with either the dG base or the AAF-unit inside the duplex (44). Melting point data of the AAF-dG-containing duplex used for crystallization and further thermodynamic studies (Fig. S4 and Table S3) show that the melting point is substantially reduced by $\Delta T_m = 12^\circ\text{C}$ ($\Delta\Delta H = 3\text{ kcal/mol}$) with no effects on ΔS ($\Delta\Delta S = 0\text{ cal/molK}$). These data prove that the AAF-dG lesion creates a thermodynamically destabilized region in the duplex that can certainly be more easily kinked. More importantly, in the Rad14 structure with the AAF-dG lesion, we observe full intercalation of both the AAF and the dG unit into the duplex (Fig. 4B and Fig. S3B). The AAF-dG takes the place of a full base pair. The dC counterbase moves in response into a flipped out conformation (no electron density for the base) toward helix $\alpha 7$. Such a position of the AAF-dG fully inside the duplex was never observed, showing the unusual situation of the molecule in the kinked DNA duplex. It is most likely that this complete intercalation of the AAF-dG unit is needed to stabilize the 70° kink using favorable π -stacking interactions. This model offers an explanation for binding of Rad14/XPA to FITC-dU containing DNA. This unit reduces the stability of the duplex by only a small amount. It may be that

XPA impose some bending force to create a kink. Once the kink is temporarily formed, it may then be stabilized by intercalation of the flat and aromatic fluorescein unit. Further crystal structures are certainly needed to clarify this question.

Our structure is also in agreement with previous detailed mutagenesis data (34, 54). The residues identified as critical contributors to the recognition of lesions by Rad14 are conserved in human XPA with the exception of Thr239, which is replaced by Lys151, Phe262, which is conservatively substituted by a tryptophane, and Gln266, which is replaced by a lysine. Although the side chain of Lys151 is somewhat longer, it would still be fully competent to fulfill the observed function of Thr239 in Rad14 (Fig. S5). The importance of the critical residues for DNA binding identified in our structure is supported by a lysine scanning mutagenesis study on human XPA where a K151E variant (corresponding to Thr239) was still able to interact with DNA but showed significantly decreased affinity to damaged DNA (54). In the same study, Lys141, Lys145, Lys179, Lys204, and Arg207 were analyzed as well. Intriguingly, the residues corresponding to Lys141 (Lys229), Lys179 (Gln266), and Arg207 (Arg294) all show impaired damaged DNA binding thus highlighting the importance of those residues for XPA function consequently validating our structural data.

The structures reported here are unable to clarify the enigmatic role of XPA in the whole NER process. However, they provide mechanistic insight into how this important NER protein is able to interact with kinked DNA structures and lesion-containing duplexes that allow sharp bending. Binding requires interrogation of the duplex with a β -hairpin structure already observed in Rad4/XPC binding to DNA (18). Here we see similarities between both proteins. In summary our structures confirm that NER proteins and as such also XPA probe the structural integrity of the duplex, which is the secret behind the broad substrate promiscuity of NER (3).

Materials and Methods

For additional information on cloning of Rad14 and XPA, protein expression and purification, DNA synthesis, crystallization, data collection, structure determination and structure refinement, protein-DNA binding studies, protein-crosslinking experiments and LC-MS analysis, DNA melting temperature measurements, and fluorescence polarization measurements please see [SI Materials and Methods](#).

ACKNOWLEDGMENTS. We thank Prof. Angelika Vollmar of the LMU Munich for access to radioactive laboratories and Jérôme Basquin and the staff of the MPI-Martinsried Crystallization Facility. Synchrotron access was supported by funds from the European Community's Seventh Framework Program (FP7/2007-2013) under grant agreement no. 226716. We also thank the beamline scientists at the SLS and ESRF for setting up the beamlines and support during data collection. We thank Dr. Markus Müller and Wolfgang Koelmel for helpful discussions, Johanna Bretzler for partly preparing and purifying XPA full length protein, and Karl-Peter Hopfner for initial crystal structure data interpretation and helpful discussions. This work was supported by Deutsche Forschungsgemeinschaft (SFB-646, SFB-749, the excellence cluster CIPSM and grant KI-562/2), Forschungszentrum FZ-82, the Fonds der Chemischen Industrie (FCI), and LMU^{excellence}. S.S. was supported by an FCI Liebig fellowship.

1. Cleaver JE (2000) Common pathways for ultraviolet skin carcinogenesis in the repair and replication defective groups of xeroderma pigmentosum. *J Dermatol Sci* 23(1): 1–11.
2. Berneburg M, Lehmann AR (2001) Xeroderma pigmentosum and related disorders: Defects in DNA repair and transcription. *Adv Genet* 43:71–102.
3. Naegeli H, Sugawara K (2011) The xeroderma pigmentosum pathway: Decision tree analysis of DNA quality. *DNA Repair (Amst)* 10(7):673–683.
4. Sancar A (1994) Mechanisms of DNA excision repair. *Science* 266(5193):1954–1956.
5. Geacintov NE, et al. (2002) Thermodynamic and structural factors in the removal of bulky DNA adducts by the nucleotide excision repair machinery. *Biopolymers* 65(3): 202–210.
6. Gillet LC, Schärer OD (2006) Molecular mechanisms of mammalian global genome nucleotide excision repair. *Chem Rev* 106(2):253–276.
7. de Laat WL, Jaspers NG, Hoeijmakers JH (1999) Molecular mechanism of nucleotide excision repair. *Genes Dev* 13(7):768–785.

8. Gunz D, Hess MT, Naegeli H (1996) Recognition of DNA adducts by human nucleotide excision repair. Evidence for a thermodynamic probing mechanism. *J Biol Chem* 271(41): 25089–25098.
9. Huang JC, Hsu DS, Kazantsev A, Sancar A (1994) Substrate spectrum of human excinuclease: Repair of abasic sites, methylated bases, mismatches, and bulky adducts. *Proc Natl Acad Sci USA* 91(25):12213–12217.
10. Jones CJ, Wood RD (1993) Preferential binding of the xeroderma pigmentosum group A complementing protein to damaged DNA. *Biochemistry* 32(45): 12096–12104.
11. Zamble DB, Mu D, Reardon JT, Sancar A, Lippard SJ (1996) Repair of cisplatin—DNA adducts by the mammalian excision nuclease. *Biochemistry* 35(31):10004–10013.
12. Kartalou M, Essigmann JM (2001) Mechanisms of resistance to cisplatin. *Mutat Res* 478(1–2):23–43.
13. Friedberg EC (2001) How nucleotide excision repair protects against cancer. *Nat Rev Cancer* 1(1):22–33.

14. Mellon I, Spivak G, Hanawalt PC (1987) Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. *Cell* 51(2): 241–249.
15. Sweder KS, Hanawalt PC (1992) Preferential repair of cyclobutane pyrimidine dimers in the transcribed strand of a gene in yeast chromosomes and plasmids is dependent on transcription. *Proc Natl Acad Sci USA* 89(22):10696–10700.
16. Tornaletti S, Hanawalt PC (1999) Effect of DNA lesions on transcription elongation. *Biochimie* 81(1–2):139–146.
17. Volker M, et al. (2001) Sequential assembly of the nucleotide excision repair factors in vivo. *Mol Cell* 8(1):213–224.
18. Min JH, Pavletich NP (2007) Recognition of DNA damage by the Rad4 nucleotide excision repair protein. *Nature* 449(7162):570–575.
19. Riedl T, Hanaoka F, Egly JM (2003) The comings and goings of nucleotide excision repair factors on damaged DNA. *EMBO J* 22(19):5293–5303.
20. Luijsterburg MS, et al. (2010) Stochastic and reversible assembly of a multiprotein DNA repair complex ensures accurate target site recognition and efficient repair. *J Cell Biol* 189(3):445–463.
21. Mathieu N, Kaczmarek N, Rüthemann P, Luch A, Naegeli H (2013) DNA quality control by a lesion sensor pocket of the xeroderma pigmentosum group D helicase subunit of TFIIH. *Curr Biol* 23(3):204–212.
22. Kuper J, Wolski SC, Michels G, Kisker C (2012) Functional and structural studies of the nucleotide excision repair helicase XPD suggest a polarity for DNA translocation. *EMBO J* 31(2):494–502.
23. Mathieu N, Kaczmarek N, Naegeli H (2010) Strand- and site-specific DNA lesion demarcation by the xeroderma pigmentosum group D helicase. *Proc Natl Acad Sci USA* 107(41):17545–17550.
24. Pugh RA, Wu CG, Spies M (2012) Regulation of translocation polarity by helicase domain 1 in SF2B helicases. *EMBO J* 31(2):503–514.
25. Sugawara K, Akagi J, Nishi R, Iwai S, Hanaoka F (2009) Two-step recognition of DNA damage for mammalian nucleotide excision repair: Directional binding of the XPC complex and DNA strand scanning. *Mol Cell* 36(4):642–653.
26. Martijn JA, Lans H, Vermeulen W, Hoeijmakers JH (2014) Understanding nucleotide excision repair and its roles in cancer and ageing. *Nat Rev Mol Cell Biol* 15(7):465–481.
27. Tapias A, et al. (2004) Ordered conformational changes in damaged DNA induced by nucleotide excision repair factors. *J Biol Chem* 279(18):19074–19083.
28. Kobayashi T, et al. (1998) Mutational analysis of a function of xeroderma pigmentosum group A (XPA) protein in strand-specific DNA repair. *Nucleic Acids Res* 26(20): 4662–4668.
29. Satokata I, Tanaka K, Yuba S, Okada Y (1992) Identification of splicing mutations of the last nucleotides of exons, a nonsense mutation, and a missense mutation of the XPAC gene as causes of group A xeroderma pigmentosum. *Mutat Res* 273(2):203–212.
30. He Z, Henriksen LA, Wold MS, Ingles CJ (1995) RPA involvement in the damage-recognition and incision steps of nucleotide excision repair. *Nature* 374(6522): 566–569.
31. Li L, Lu X, Peterson CA, Legerski RJ (1995) An interaction between the DNA repair factor XPA and replication protein A appears essential for nucleotide excision repair. *Mol Cell Biol* 15(10):5396–5402.
32. Park CH, Mu D, Reardon JT, Sancar A (1995) The general transcription-repair factor TFIIH is recruited to the excision repair complex by the XPA protein independent of the TFIIIE transcription factor. *J Biol Chem* 270(9):4896–4902.
33. Tripsianes K, et al. (2007) Analysis of the XPA and ssDNA-binding surfaces on the central domain of human ERCC1 reveals evidence for subfunctionalization. *Nucleic Acids Res* 35(17):5789–5798.
34. Camenisch U, Dip R, Schumacher SB, Schuler B, Naegeli H (2006) Recognition of helical kinks by xeroderma pigmentosum group A protein triggers DNA excision repair. *Nat Struct Mol Biol* 13(3):278–284.
35. Missura M, et al. (2001) Double-check probing of DNA bending and unwinding by XPA-RPA: An architectural function in DNA repair. *EMBO J* 20(13):3554–3564.
36. Evans E, Moggs JG, Hwang JR, Egly JM, Wood RD (1997) Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. *EMBO J* 16(21):6559–6573.
37. Liu Y, et al. (2005) Cooperative interaction of human XPA stabilizes and enhances specific binding of XPA to DNA damage. *Biochemistry* 44(19):7361–7368.
38. Yang ZG, Liu Y, Mao LY, Zhang JT, Zou Y (2002) Dimerization of human XPA and formation of XPA2-RPA protein complex. *Biochemistry* 41(43):13012–13020.
39. Kuraoka I, et al. (1996) Identification of a damaged-DNA binding domain of the XPA protein. *Mutat Res* 362(1):87–95.
40. Ikegami T, et al. (1998) Solution structure of the DNA- and RPA-binding domain of the human repair factor XPA. *Nat Struct Biol* 5(8):701–706.
41. Schweizer U, Hey T, Lipps G, Krauss G (1999) Photocrosslinking locates a binding site for the large subunit of human replication protein A to the damaged strand of cis-platin-modified DNA. *Nucleic Acids Res* 27(15):3183–3189.
42. Buchko GW, Ni S, Thrall BD, Kennedy MA (1998) Structural features of the minimal DNA binding domain (M98-F219) of human nucleotide excision repair protein XPA. *Nucleic Acids Res* 26(11):2779–2788.
43. Sugitani N, Shell SE, Soss SE, Chazin WJ (2014) Redefining the DNA-binding domain of human XPA. *J Am Chem Soc* 136(31):10830–10833.
44. Lukin M, de Los Santos C (2006) NMR structures of damaged DNA. *Chem Rev* 106(2): 607–686.
45. Krasikova YS, et al. (2013) Comparative analysis of interaction of human and yeast DNA damage recognition complexes with damaged DNA in nucleotide excision repair. *J Biol Chem* 288(15):10936–10947.
46. Sugawara K, et al. (1998) Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. *Mol Cell* 2(2):223–232.
47. Sugawara K, et al. (2001) A multistep damage recognition mechanism for global genomic nucleotide excision repair. *Genes Dev* 15(5):507–521.
48. Hey T, et al. (2002) The XPC-HR23B complex displays high affinity and specificity for damaged DNA in a true-equilibrium fluorescence assay. *Biochemistry* 41(21):6583–6587.
49. Fujiwara Y, et al. (1999) Characterization of DNA recognition by the human UV-damaged DNA-binding protein. *J Biol Chem* 274(28):20027–20033.
50. Scrima A, et al. (2008) Structural basis of UV DNA-damage recognition by the DDB1-DDB2 complex. *Cell* 135(7):1213–1223.
51. Li L, Peterson CA, Lu X, Legerski RJ (1995) Mutations in XPA that prevent association with ERCC1 are defective in nucleotide excision repair. *Mol Cell Biol* 15(4):1993–1998.
52. Buschta-Hedayat N, Buterin T, Hess MT, Missura M, Naegeli H (1999) Recognition of nonhybridizing base pairs during nucleotide excision repair of DNA. *Proc Natl Acad Sci USA* 96(11):6090–6095.
53. Buterin T, Meyer C, Giese B, Naegeli H (2005) DNA quality control by conformational readout on the undamaged strand of the double helix. *Chem Biol* 12(8):913–922.
54. Camenisch U, Dip R, Vitanescu M, Naegeli H (2007) Xeroderma pigmentosum complementation group A protein is driven to nucleotide excision repair sites by the electrostatic potential of distorted DNA. *DNA Repair (Amst)* 6(12):1819–1828.
55. Wang Z, Rizzo CJ (2001) Synthesis of the C8-deoxyguanosine adduct of the food mutagen IQ. *Org Lett* 3(4):565–568.
56. Gillet LC, Schärer OD (2002) Preparation of C8-amine and acetylamino adducts of 2'-deoxyguanosine suitably protected for DNA synthesis. *Org Lett* 4(24):4205–4208.
57. Gillet LC, Alzeer J, Schärer OD (2005) Site-specific incorporation of N-(deoxyguanosin-8-yl)-2-acetylaminofluorene (dG-AAF) into oligonucleotides using modified 'ultra-mild' DNA synthesis. *Nucleic Acids Res* 33(6):1961–1969.
58. Wei M, Cohen SM, Silverman AP, Lippard SJ (2001) Effects of spectator ligands on the specific recognition of intrastrand platinum-DNA cross-links by high mobility group box and TATA-binding proteins. *J Biol Chem* 276(42):38774–38780.
59. Allawi HT, SantaLucia J, Jr (1997) Thermodynamics and NMR of internal G.T mismatches in DNA. *Biochemistry* 36(34):10581–10594.
60. Zaliznyak T, Bonala R, Johnson F, de Los Santos C (2006) Structure and stability of duplex DNA containing the 3-(deoxyguanosin-N2-yl)-2-acetylaminofluorene (dG(N2)-AAF) lesion: A bulky adduct that persists in cellular DNA. *Chem Res Toxicol* 19(6): 745–752.
61. LeMaster DM, Richards FM (1985) ¹H-¹⁵N heteronuclear NMR studies of Escherichia coli thioredoxin in samples isotopically labeled by residue type. *Biochemistry* 24(25): 7263–7268.
62. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50(Pt 5):760–763.
63. Kabsch W (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):133–144.
64. Vonrhein C, et al. (2011) Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):293–302.
65. Terwilliger TC, et al. (2009) Decision-making in structure solution using Bayesian estimates of map quality: The PHENIX AutoSol wizard. *Acta Crystallogr D Biol Crystallogr* 65(Pt 6):582–601.
66. Zwart PH, et al. (2008) Automated structure solution with the PHENIX suite. *Methods Mol Biol* 426:419–435.
67. Terwilliger TC, et al. (2008) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr* 64(Pt 1):61–69.
68. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213–221.
69. McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr* 61(Pt 4):458–464.
70. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Cryst* 40(Pt 4):658–674.
71. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501.
72. Murshudov GN, et al. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):355–367.
73. Winn MD, Murshudov GN, Papiz MZ (2003) Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol* 374:300–321.
74. Painter J, Merritt EA (2006) TLSMD web server for the generation of multi-group TLS models. *J Appl Cryst* 39(1):109–111.
75. Shevchenko A, Wilm M, Vorm O, Mann M (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem* 68(5):850–858.
76. Sievers F, et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7:539.
77. Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* 292(2):195–202.

Supporting Information

Koch et al. 10.1073/pnas.1508509112

SI Materials and Methods

DNA Preparation. The AAF-dG phosphoramidite containing an isopropylphenoxyacetyl group at the N²-position for solid phase DNA synthesis was prepared as published (55, 56) and incorporated into DNA using ultra mild conditions for DNA synthesis (57). Oligonucleotide synthesis was performed on an ABI 394 Nucleic Acid Synthesis System (Life Technologies). Phosphoramidites for dA, dC, dG, dT, and CPG carriers were obtained from Glen Research or Link Technologies. The coupling time for the modified phosphoramidites was extended to 2 × 10 min. For the generation of the 1,2-GG cisplatin containing oligonucleotides, a modified protocol was used as published (58). Cisplatin was activated by treating the platinum salt with two equivalents of silver nitrate in an aqueous solution at a final concentration of 5 mM/10 mM. The mixture was protected from light and incubated at 18 °C for 15 h. Precipitated silver chloride was removed by centrifugation. The oligonucleotide 5'-TCTCTACGGTTCAT-CAC-3' was platinated with 3.0 equivalents of the activated platinum complex in a solution containing 100 mM NaClO₄ at 37 °C for 4 h. The platinated oligonucleotide was purified by reverse phase HPL chromatography (Macherey-Nagel, Nucleosil 100-7 C18, 10 × 250 mm, 0.1 M triethylammonium acetate, linear gradient from 8% to 20% or 0–40% acetonitrile in 45 min). The purified strands were desalted using C18 cartridges (Waters, Sep-Pak C18 Classic Cartridge). The identity of the oligonucleotides were confirmed by matrix-assisted laser-desorption time-of-flight mass spectrometry and the purity determined to be >98% by capillary electrophoresis.

Melting Temperatures and Thermodynamic Parameters. The duplexes of ODN6 (at concentrations between 1 and 7 μM; see Table S1) were dissolved in 200 μL of 10 mM Tris-HCl, pH 7.4, containing 150 mM NaCl. The samples were heated from 10 to 80 °C or cooled from 80 to 10 °C with a rate of 0.5 °C per min. Three melting profiles per sample concentration were performed with six different concentrations for each duplex. The enthalpy (ΔH°) and entropy (ΔS°) were calculated using a published protocol (59, 60). Van't Hoff plot's of T_m^{-1} versus $\ln(C_t)$ to fit to $T_m^{-1} = R/\Delta H^\circ \ln C_t/4 + \Delta S^\circ/\Delta H^\circ$, in which T_m is a melting transition point in K, C_t is the total duplex concentration, and R is the universal gas constant (1.978 cal/Kmol); $\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ$. Reported enthalpy and entropy values are the average of the determinations, and error estimates represent the SD of the data. The Gibbs free energy (ΔG°) of duplex formation was then calculated from these values with $T = 298\text{K}$.

Cloning of Rad14. The sequence coding for the minimal DNA binding domain of the *S. cerevisiae* Rad14 protein (Rad14_{188–306}) was amplified by PCR from genomic *S. cerevisiae* DNA using the AccuPrime Pfx DNA polymerase (Invitrogen) with the primers: forward: 5'-Pho-AATGGCGCCGAAATGTATTGAATGT-3' and reverse: 5'-Pho-TCCC-GTATTTTTTCTCCCTTCTGTG-3'. The resulting PCR product was cloned via the StarGate method (IBA) into the pPSG-IBA3 with a C-terminal Strep-tagII. A truncated version of the minimal DNA binding domain (Rad14_{t188–302}) lacking 4 amino acids on its C-terminal end was amplified by PCR using High Fidelity Phusion DNA polymerase (Finnzymes) and the following primers: forward: 5'-AGCGGCTCTTCAAT-GCTGGAAGTTCTGTTCCAGGGGCCAATGGCGCCGAA-ATGTATTGAATGTC-3' and reverse: 5'-AGCGGCTCTTCT-CCCCCTTCTGTGAGCCTTT-CCTTC-3'. The Rad14_{tF262A} point mutation was generated through site-directed mutagenesis using

the Q5 high fidelity polymerase (NEB) and the following primers: forward 5'-TTCGGGGACAGCGGCAAGAATGCAAC-3' and reverse 5'-TGAGGGTTTCGGCTTT-TCTAG-3'. Rad14_t and Rad14_{tF262A} were expressed with an N-terminal His-tag. A truncated version of the full length protein Rad14 (Rad14) lacking 9 amino acids on its N-terminal end was amplified by PCR using High Fidelity AccuPrime Pfx DNA polymerase (Invitrogen) and the following primers: forward: 5'-Pho-AATGGAGGCTAACA-GGAAATTAGCAATAG-3' and reverse: 5'-Pho-TCCCAATGT-CAATTTCTTCAGTTTCTAGCC-3'. The resulting PCR product was cloned via the StarGate method (IBA) into the pPSG-IBA3 with a C-terminal Strep-tagII.

Cloning of XPA. The sequence coding for the minimal DNA binding domain of the human XPA protein (XPA_{98–219}) was amplified by PCR from genomic DNA using the Phusion DNA Polymerase (NEB) with the following primers: forward: 5'-Pho-AATGGAA-TTTGATTATGTAATATGCG-3' and reverse: 5'-Pho-TCCCA-AATTTCTTCTGTTTCATTTTTTCTC-3'. The resulting PCR product was cloned via the StarGate method (IBA) into the pPSG-IBA3 with a C-terminal Strep-tagII. The sequence coding for the full length human protein (XPA_{fl}) was amplified by PCR from genomic DNA, using the Phusion DNA Polymerase (NEB) with the following primers: forward: 5'-Pho-AATGGCGGCGGCCG-ACGG-3' and reverse: 5'-Pho-TCCCCATTTTTTTCATATGTC-AGTTCATGGCC-3'. The resulting PCR product was cloned via the StarGate method (IBA) into the pPSG-IBA35 with an N-terminal His-tag.

Protein Expression and Purification. For protein expression, the constructs were transformed into *Escherichia coli* BL21 Star (DE3) (Invitrogen) cells and for the generation of selenomethionine derivatized protein *E. coli* B834(DE3)*pLysS* cells were used. Protein expression was induced by addition of isopropyl-β-D-galactoside (0.5 mM) at 25 °C (Rad14) or 28 °C (XPA) for 3 h, and the cells were grown either in Luria-Bertani or LeMaster (61) medium, respectively, supplemented with ZnCl₂ (10 μM). All purification steps were carried out at 4 °C, and the purification was monitored by SDS/PAGE. For purification of Rad14-His, the cells were resuspended in 100 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM DTT, 10% glycerol, and 10 μM ZnCl₂. For the purification using the strep-tagged Rad14 and XPA, cells were resuspended in 100 mM Tris-HCl pH 8.0, 150 mM NaCl, and 1 mM EDTA. Buffers were supplemented with the Complete protease inhibitor mix (Roche). Cells were lysed in a French press, and the cell debris was removed by centrifugation before the supernatant was applied to the NiNTA column (GE Healthcare) or to the Strep-Tactin column (IBA), respectively. Protein elution was performed either with 500 mM imidazole (His-tag) or 2.5 mM desthiobiotin (Strep-tagII). Following tag-mediated affinity purification, the fractions were pooled and concentrated in MonoQ buffer A (100 mM Tris pH 7.5, 100 mM NaCl, 5% glycerol, 5 mM DTT, and 10 μM ZnCl₂), and the protein was further purified by ion exchange (Q Sepharose from GE Healthcare) chromatography. The protein was eluted with a gradient of 5 column volumes with MonoQ buffer A containing 800 mM NaCl. Excess salt was removed by repeated concentration and dilution in 50 mM Hepes pH 8.0, 50 mM MgCl₂, 5% glycerol, 5 mM DTT, and 10 μM ZnCl₂ (crystallization buffer). In a final step, the protein was purified by size exclusion chromatography using a Superdex 75 column (GE Healthcare) in crystallization buffer. The correct mass of the protein was confirmed by SDS/PAGE and MALDI-TOF analysis. For

purification of the selenomethionine derivatized protein, 10 mM DTT was added to each buffer.

Protein Expression and Purification for the Full-Length Proteins. For Rad14fl protein expression, the constructs (pPSG-IBA3, C-terminal Strep-tagII) were transformed into *E. coli* BL21 Star (DE3) (Invitrogen) cells; for the generation of XPAfl protein (pPSG-IBA35, N-terminal His-tag), *E. coli* BL21(DE3)pLysS cells were used. Protein expression was induced by addition of isopropyl- β -D-galactoside (0.5 mM) at 25 °C (Rad14) or 37 °C (XPA) for 4 h and the cells were grown in Luria-Bertani (7) medium, respectively, supplemented with ZnCl₂ (10 μ M). All purification steps were carried out at 4 °C, and the purification was monitored by SDS/PAGE. For purification of XPA-His, the cells were resuspended in 25 mM Tris-HCl pH 7.5, 100 mM NaCl, 1 mM DTT, 10% glycerol, and 10 μ M ZnCl₂. For the purification of Rad14fl, cells were resuspended in 100 mM Tris-HCl pH 8.0, 800 mM NaCl and 1 mM EDTA, 5 mM DTT, 5% glycerol, and 10 μ M ZnCl₂. Buffers were supplemented with the Complete protease inhibitor mix (Roche). Cells were lysed by sonication and the cell debris was removed by centrifugation before the supernatant was applied to the NiNTA column (GE Healthcare) or to the StrepTactin column (GE Healthcare), respectively. Protein elution was performed either with 500 mM imidazole (His-tag) or 2.5 mM desthiobiotin (Strep-tagII). Following tag-mediated affinity purification, the fractions were pooled and concentrated in Heparin buffer A for Rad14 (100 mM Tris pH 7.5, 100 mM NaCl, 5% glycerol, 5 mM DTT (DTT) and 10 μ M ZnCl₂) and in Heparin buffer C for XPA (25 mM Tris pH 7.5, 50 mM NaCl, 10% glycerol, 1 mM DTT and 10 μ M ZnCl₂), and the protein was further purified by a heparin column (GE Healthcare) chromatography. Rad14fl was eluted with a gradient of 10 column volumes with Heparin buffer A containing 600 mM NaCl. Excess salt was removed by repeated concentration and dilution in Heparin buffer A. XPAfl was eluted with a gradient of 20 column volumes with Heparin buffer C containing 1 M NaCl. Excess salt was removed by repeated concentration and dilution in 50 mM Hepes pH 8.0, 50 mM MgCl₂, 5% glycerol, 5 mM DTT, and 10 μ M ZnCl₂ (crystallization buffer). In a final step, XPAfl was purified by size exclusion chromatography using a Superdex 75 column (GE Healthcare) in crystallization buffer. The correct mass of both proteins were confirmed by SDS/PAGE and MALDI-TOF analysis.

Crystallization, Data Collection, Structure Determination, and Structure Refinement. For cocrystallization experiments AAF-dG containing 15mer DNA (ODN6, see Table S1) was annealed in crystallization buffer to its counter strand without (5'-GTG ATG ACG TAG AGC-3') or with 5-iododeoxyuridine (U*) (5'-GU*G AU*G ACG U*AG AGC-3') (Metabion). For cocrystallization experiments with the cisplatin adduct a 16mer DNA strand (ODN7, see Table S1) was annealed to its counter strand in crystallization buffer. Before crystallization protein and DNA were mixed in a molar ratio of 1:1 protein:DNA (Strep-tagII) and 2:1 protein:DNA (His-tag) and incubated for 30 min at 4 °C. Initially, crystals for the selenomethionine labeled Strep-tagged protein in complex with AAF-containing DNA were obtained in 0.05 M Tris-HCl, pH 7, 0.1 M calcium acetate and 16–18% PEG8000 at 4 °C. His-tagged Rad14t protein in complex with AAF-dG and U*- or 1,2-GG cisplatin containing DNA was crystallized in 0.18 M ammonium nitrate and 40% 2-methyl-1,3-propanediol (MPD) at 4 °C. Crystals were frozen in artificial mother liquor supplemented with 15% (vol/vol) ethylene glycol (SeMet-Rad14) or directly from the well (Rad14t) and stored in liquid nitrogen until data collection. Single anomalous dispersion (SAD) data on the SeMet-labeled Rad14-AAF complex crystals were collected at the microdiffractometer at the PXI beamline (Swiss Light source, Villigen, Switzerland). The crystals belonged to the space group P2₁, with unit cell dimensions of $a = 66.4$ Å, $b = 51.7$ Å,

$c = 68.8$ Å and $\beta = 109.6^\circ$ and diffracted X-rays anisotropically to about 2.8 Å spacing. For structure determination, data collected on four different crystals were combined using AUTOPROC (62–64) and experimental phases were determined using PHENIX Autosol (65, 66). Initial model building was carried out with PHENIX Autobuild (67, 68). Iodine SAD data were collected at beamline ID29 of the European Synchrotron Radiation Facility (ESRF), diffracted X-rays to 1.8 Å and belonged to the tetragonal space group P4₁, with unit cell dimensions of $a = 53.72$ Å, $b = 53.72$ Å, and $c = 130.55$ Å. Crystals of Rad14t in complex with the 1,2-GG cisplatin DNA diffracted X-rays to 2.8 Å at the microfocus beamline ID23-2 (ESRF) and also belonged to the space group P4₁ with comparable unit cell constants as the AAF-dG complex. The structures were solved by molecular replacement (PHASER; refs. 69 and 70) using the coordinates of the Rad14-SeMet protein. Manual building of the model was performed with COOT (71) and structure refinement in Refmac5 (72) and Phoenix (68) using TLS (73, 74). Positions of the iodine and platinum atoms were determined and their occupancy estimated from the anomalous difference data (63). Structural figures were prepared with PyMOL (Delano Scientific).

Electromobility Shift Assays and Determination of the Binding Constant.

First, the counterstrand oligonucleotides were 5'-labeled with [γ -³²P]ATP (3,000 Ci/mmol) using T4 polynucleotide kinase. Then, the damaged and undamaged oligonucleotides were mixed with their labeled complementary strands in equal molar amounts and annealed to form duplex DNA. The formation of the duplex DNA was checked by native TBE polyacrylamide gel electrophoresis. A standard DNA binding reaction mixture (10 μ L) contained 33 fmol of the labeled duplex DNA and varying concentrations of purified Rad14 or XPA in binding buffer [25 mM Hepes-KOH pH 8.3, 100 mM KCl, 1 mM EDTA, 1 mM DTT, 45 μ g/mL BSA and 10% (vol/vol) glycerol]. After incubation at 30 °C for 30 min, 2 μ L of loading dye (250 mM Tris-HCl pH 7.5, 40% glycerol, 0.2% bromophenolblue) were added and electrophoresis was performed on a 4% nondenaturing polyacrylamide TBE gel at 4 °C, which was prerun for 30 min. Band shifts were visualized by autoradiography (Storm TM860, GE). Each gel-shift reaction and gel was performed three times. The binding constant was determined by a competitive binding experiment. The concentration of a radio labeled lesion-containing DNA was kept constant while increasing amounts of competing unlabeled DNA were added. The binding assay solutions contained 33 fmol ³²P-labeled, lesion-containing DNA, 23 pmol purified Rad14 and varying amounts of lesion containing unlabeled DNA (Fig. S1H).

Fluorescence Polarization Measurements. The DNA binding specificity and cooperativity of Rad14_{188–302} was determined in a fluorescence polarization experiment, using internal fluorescein-labeled DNA (30mer ODN10 and 15mer 11, see Table S1) substrate. The experiments were performed at 25 °C on a GeniusPro fluorescence photometer (Tecan). The excitation wavelength was set at 485 nm, and the emission wavelength was set at 535 nm. Measurements with Rad14t and Rad14t_{F262A} were carried out in binding buffer (total volume of 100 μ L) by incubating increasing amounts of protein with double stranded DNA containing an internal FITC-dU (5 nM) for 30 min at 25 °C. Measurements were carried out in 10 \times 4-mm cuvettes at ambient temperature. Ten data points were taken per titration point. Each experiment was repeated three times.

Protein Crosslinking and LC-MS Analysis. For the crosslinking experiments Rad14t was incubated for 30 min at 4 °C with a 19mer DNA duplex (ODN8, see Table S1) or a 37mer DNA duplex (ODN9) containing AAF-dG at its central position. Increasing amounts of Bis(NHS)PEG₅ crosslinker were added to the protein-DNA complex (1:1–1:200) and the mixture was incubated

for 2 h at 4 °C. After adding 2 μ L of loading dye a SDS PAGE was performed. The protein band of interest was excised using a scalpel and subsequently in-gel digested as described by Shevchenko et al. (75). Trypsin and chymotrypsin were used for protein digestion. The samples were analyzed using an UltiMate 3000 nano liquid chromatography system (Dionex, Fisher Scientific) coupled to an LTQ-Orbitrap XL (Fisher Scientific). Ten microliters of the protein digest were injected, the samples desalted, concentrated on a μ -precolumn cartridge (PepMap100, C18, 5 μ M, 100 \AA , size 300 μm i.d. \times 5 mm) and further processed on a custom made analytical column (ReproSil-Pur, C18, 3 μ M, 120 \AA , packed into a 75 μm i.d. \times 150 mm and 8 μm picotip emitter). An 85-min multistep analytical separation was performed at a flow rate of 300 nL/min. In the first 50 min, a linear gradient was ramped up from 5% solvent B (acetonitrile containing 3% DMSO and 0.1% formic acid) and 95% solvent A (water containing 3% DMSO and 0.1% formic acid) to 95% solvent B. This level was held for 5 min and then ramped down again to 5% solvent B within

2 min. Mass spectrometric analyses were performed starting with a full mass scan in the mass range between m/z 300 and 1650. This survey scan was followed by three MS/MS scans using the FTMS mass analyzer and normalized collision energy of 70 in the HCD cell and three additional scans using the ion trap mass analyzer and normalized collision energy of 35. The Thermo Proteome Discoverer 1.1 software (Fisher Scientific) was used for protein identification. The Sequest search engine was used in combination with a uniprot database of the expression system (*E. coli*, complete proteome, download 12.2013) and the fasta file of the protein of interest "DNA repair protein RAD14" (*S. cerevisiae*, download 12.2013). As limit of detection a ratio of a threefold signal over the noise filter was applied. A maximum of three missed cleavage sites was allowed. The mass tolerances were 30 ppm for the precursor mass and 0.8 Da for the fragment ion mass. Dynamic modifications were cation: Na (Asp, Glu), oxidation (Met) as well as possible modifications of the crosslinker: Bis(NHS)PEG₅ (Lys).

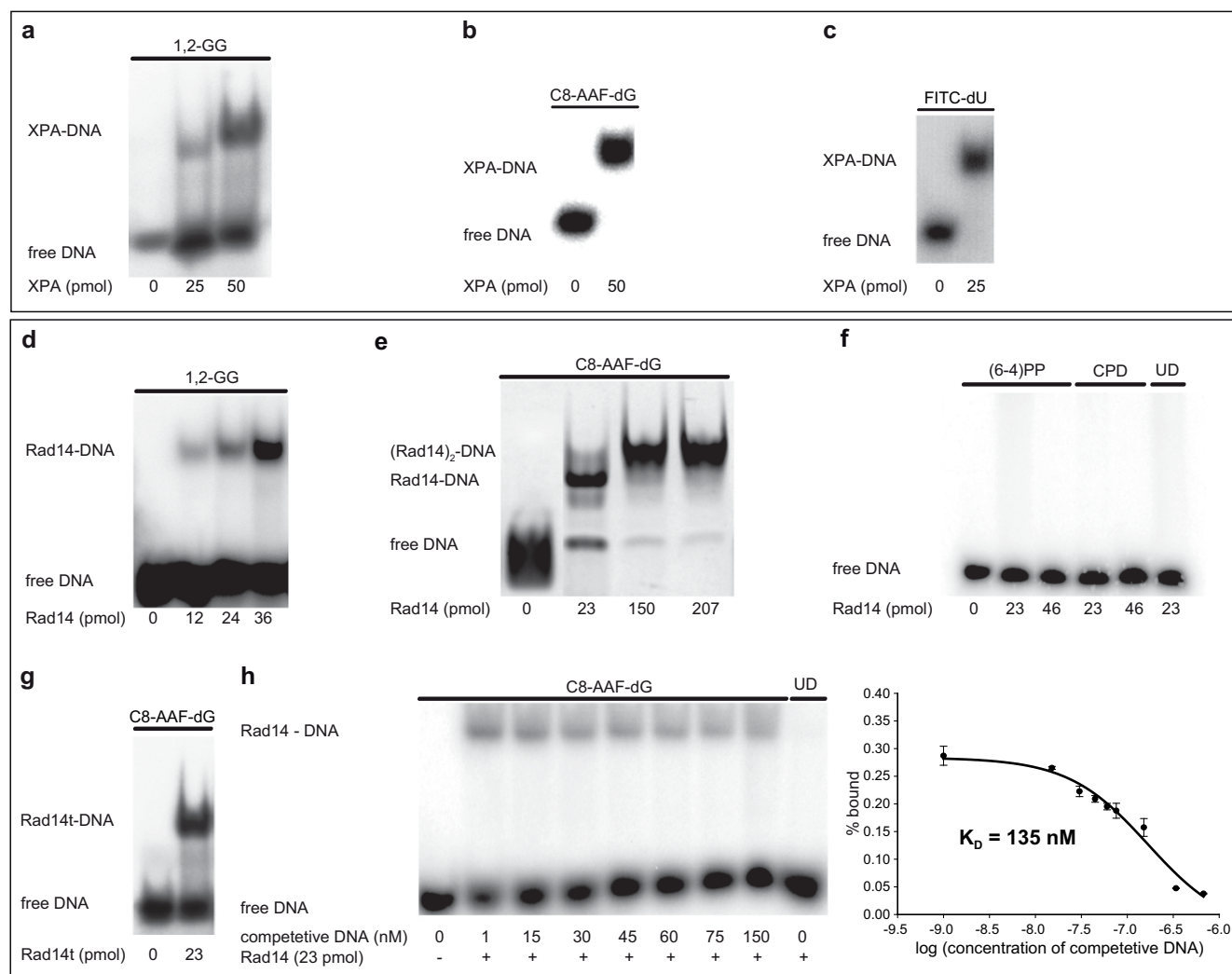
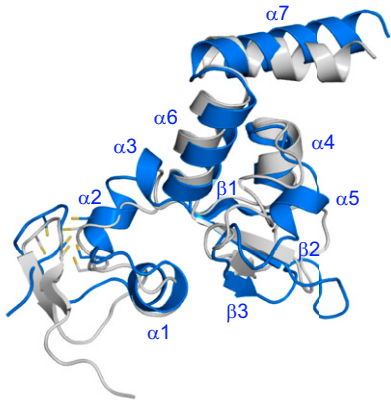


Fig. S1. Binding assays of XPA and Rad14 with different DNA lesions. The indicated amounts of protein were incubated with DNA duplexes (33 fmol) containing the lesions. As control, undamaged dsDNA (UD) was used. Electrophoretic mobility shift assay (EMSA) of XPA with 1,2-GG (A, ODN1 see Table S1), AAF-dG (B, ODN2), and FITC-dU (C, ODN3). EMSA of Rad14 1,2-GG (D, ODN1), AAF-dG (E, ODN2) and UV-lesions [(6, 4)PP and CPD; F, ODN4]. (G) EMSA of Rad14₁₈₈₋₃₀₂ (Rad14t) with AAF-dG containing DNA (ODN2). (H) Competitive EMSA for the determination of the binding constant of Rad14 to AAF-dG (ODN2). Indicated amounts of Rad14 were incubated with DNA containing AAF-dG (lanes 2–8) and unmodified DNA (lane 9). Competitive DNA was added to the mixture as indicated. The amount of the labeled DNA was counted and assigned as the percentage of bound DNA in the y axis. The x axis shows the logarithmic concentration of the added competition DNA. The curve was fitted as a one-site competition experiment. A binding constant of $K_{\text{aff}} = 135 \text{ nM}$ ($\pm 10 \text{ nM}$ SEM) was determined for the affinity of Rad14 to DNA duplexes containing this bulky adduct.

a



b

XPA_MOUSE	1	MATA EE KQTS PE PV-AADE PA Q-----L PA AVRASVER----
XPA_HUMAN	1	--MAAADGAL PE AA-ALE Q PAE-----L PA SVRASIER----
Rad14_Aspergillus	1	--MVTE EQ SAPPPERGMST PG DLKRNP LT PEQLRRIENN RQ K---AI AI RE--Q RE ---
Rad14_Saccharomyces	1	--MT PE Q AK LEAN RK LA IE RL RK GI LSS Q LN RI ES RNEPL KTR PLAVTSGSN R DDNA
psipred output	1	CCHHHHHHHHHHHHHHHHHHHHHCCCCCHHHHHHHHHHHCCCCCCCCCHHHHHHHHHCCCCHH
XPA_MOUSE	33	----- KRQ RALM-----LR-QAR L AAR PY PA--AAATGGV AS
XPA_HUMAN	31	----- KRQ RALM-----LR-QAR L AAR PY SATAAAATGG MA N
Rad14_Aspergillus	50	--AEQARA HN AD PS STKG V KRTYSS M -TASDPPAT LR DA AS TDN R PL DG IK PA RNF T K FV
Rad14_Saccharomyces	59	AAAVHVP NH NG Q PSALANTNT NT TS LY GSGVVD GSK RDAS VL DK RPT DR IR PS IR KQDY I
psipred output	60	HHCCCCCCCCCCCCCCCCCCCCCCCCCHHCCCCCCCCCHHCCCCCCCCCCCCCCCCCCCCCE
XPA_MOUSE	62	VKAAP K MIDTKGGFILE EEEE -EK HE -----IGN I V HE P
XPA_HUMAN	62	VKAAP K IIDTKGGFILE EEEE EE EQ K-----IG K V VH Q P
Rad14_Aspergillus	107	EYDFS K ITDT TK GG FL TEED DR Y NK ALHVRDE KA --EQ K PAH MT Q K E WE R Q Q L L KN LHR DR
Rad14_Saccharomyces	119	EYDFAT M Q N LNGGY I NP KD KL PN SDF TD DQ E FESE F GSK KQ KT LQ D WK KE Q LER KM LY EN
psipred output	120	EECHHHHCECCCCCCCCCCCCCCCCCCCCCCCCCHHCHHHHHHHHHHHHHHHHHHCCC
XPA_MOUSE	94	GPVM-----EF I YT IC EE CG -KEF M DSY LM NHFD LP TC DS CR DA -DDKH KLI T KTE
XPA_HUMAN	95	GPVM-----EF I Y VI CE EC G-KEF M DSY LM NHFD LP TC DN CR DA -DDKH KLI T KTE
Rad14_Aspergillus	165	AGPF E PGLSVLDERT Q KT CR EE G -S LE ID WK WEEALRCC V CH AC K DK FP E K YS LLT KTE
Rad14_Saccharomyces	179	APP PE -----HIS K AP K IE CH INIEMDP VL HD VF L Q V CK Q CS KEHP E K Y ALLT KTE
psipred output	179	CCCCCCHHHCHHHHHHHHHHHHHHHHHHHHHCCCCCHHCCCCCHHHCCCCCCEECHHH
XPA_MOUSE	143	AKQ E Y LI L K OC D L E K RE PA L R FL V KN PH SQ WGDM K LY L K LQ V V K R AL--EVWGS Q E AL E
XPA_HUMAN	144	AKQ E Y LI L K OC D L E K RE PP L K F IV KN PH SQ WGDM K LY L K LQ IV K RS L --EVWGS Q E AL E
Rad14_Aspergillus	224	AR E D YL IT N PE LQ DE-E LL PH LE R NP H K ST W NN M LY L R YQ VEE YA FS AK KWGS PE AL D
Rad14_Saccharomyces	232	CK E D Y FI T DP EL NDE-D LF H RL E K EN PH SG T F AR M Q L F V R CE VE AF AF --KKWGS E GL D
psipred output	233	HHHHHCCCCCHHHCCC CCCCCCCCCCCCCCCCCCEEEEEHHHHHHHHH HHHCHHHHHH
XPA_MOUSE	201	DA K F V R EN RE K M K Q K F D K V K V GLDDAVKILLDD K PL HR M-AP-----WH N R A LE K RS
XPA_HUMAN	202	E A K F V R ENRE K M K Q K F D K V K EL RR AV -----RS-SV-----W K RE T I V H Q H
Rad14_Aspergillus	283	AE F E R ENE K KRR EV K F KS L ED L K K R TR VD AY RRSR Q GAT G GN F GDD L GG G GR H V HQ W
Rad14_Saccharomyces	289	EE W Q R EE G KA HR RE K K Y E K I K EM RL K TR AQ EY T N RL RE KK-----HG K A H I H HF
psipred output	290	HHCCCCCCCC
XPA_MOUSE	255	SVL TR TTTT KQ IS FI YSIN TC DC CA K-----
XPA_HUMAN	245	EYGP---E N LEDD MY R K T CT MC G HE L TY E K M
Rad14_Aspergillus	343	GR SV ---ED P KT GI GV K CV DC GM EV EE LE F
Rad14_Saccharomyces	340	SD P V D GGI DE DG YQ I Q RR R CT DC GL ET EE IDI
psipred output	341	CCCCCCCCCCCCCEEEEEEECCCCCEEEEEEC

Fig. S2. Comparison of the human XPA and its yeast homolog Rad14. (A) Structural superposition of the yeast Rad14 DNA-binding domain with the NMR structure (Protein Data Bank ID code 1XPA) of its human homolog (Sequence identity 26.5%, rmsd: 2.2 Å). (B) Sequence alignment of Rad14/XPA homologs. Sequences were aligned with Clustal Omega (76). The secondary structure prediction of the *Saccharomyces* enzyme was performed with psipred and is shown in the bottom line of the alignment. H represents α -helices, C coils, and E β -strands (77). The boxed area represents the structure shown in the manuscript. The color scheme follows that of Fig. 2. The cysteine residues that build the zinc finger are boxed in yellow.

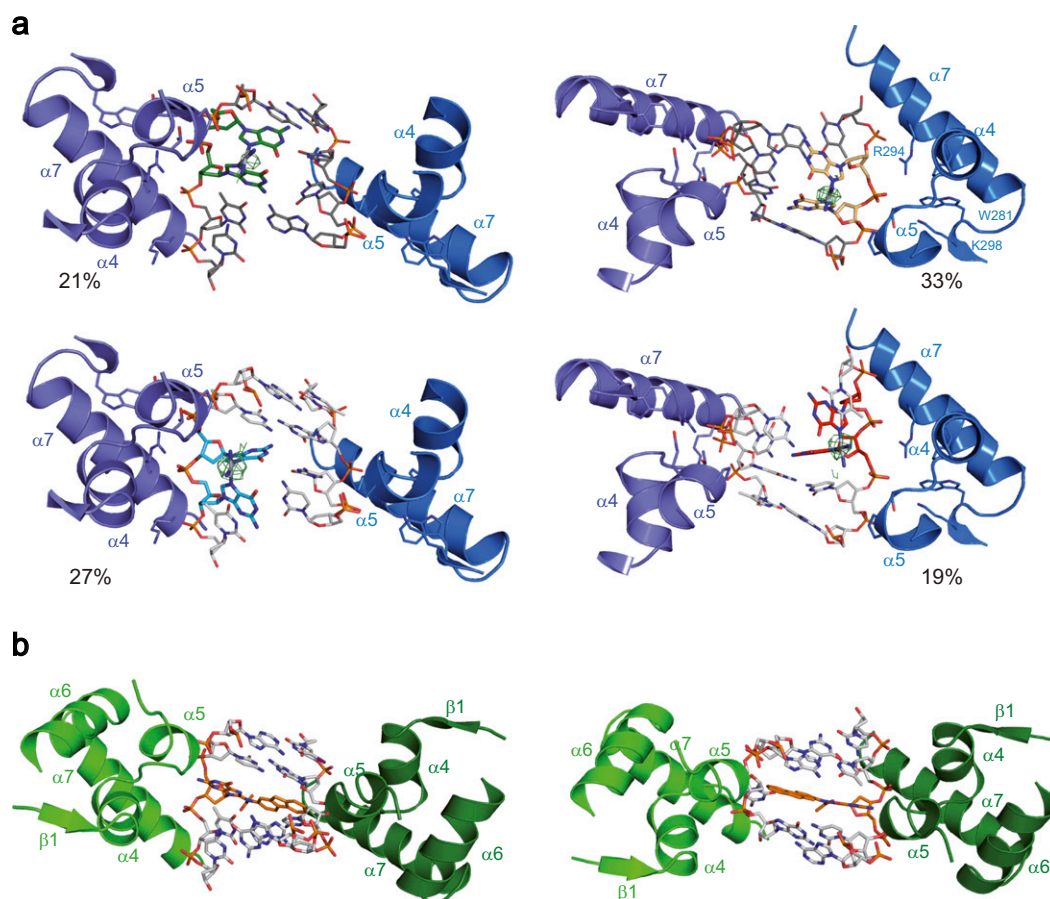


Fig. S3. Alternative orientations of the DNA bound by two Rad14 proteins. (A) The dinucleotide 1,2-GG lesion occupies four alternative positions in the 13mer duplex wedged between the β -hairpins of the two Rad14 proteins. The occupancy was estimated from the anomalous signal of the platinum atoms. (B) The AAF-dG lesion takes up two opposite positions in the crystal structure.

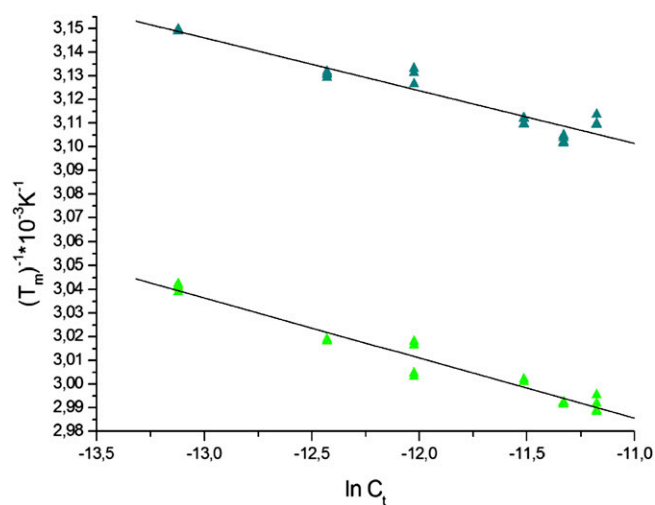


Fig. S4. Van't Hoff plots of T_m^{-1} vs. $\ln(C_t/4)$ derived from the melting curves for the control (undamaged DNA, green) and AAF-dG 15mer duplexes (blue) at concentrations between 2 μM and 7 μM . C_t represents the total oligonucleotide concentration.

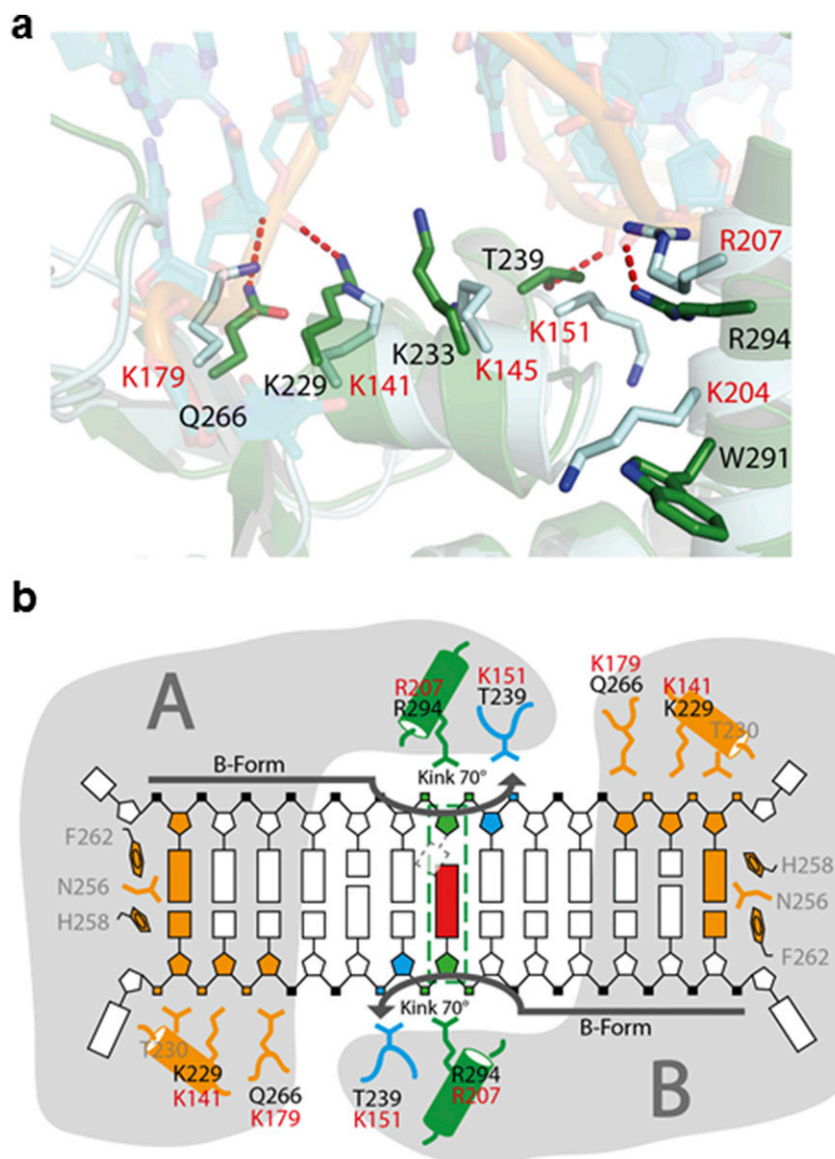


Fig. S5. Comparison of Rad14 bound to DNA and human XPA. (A) Superposition of a Rad14 monomer (green) bound to DNA (cyan) with XPA (light blue). Residues that have been implicated in DNA binding for XPA are shown in stick mode. The corresponding residues in Rad14 are also depicted. Human XPA residue numbers are shown in red; Rad14 residue numbers are shown in black. (B) Schematic representation of the Rad14 binding mode to DNA using the scheme introduced in Fig. 2D. Human XPA residues involved in DNA binding are depicted in red.

Table S1. Sequences of the oligonucleotides (ODN) used in the different experiments

ODN	Oligonucleotide sequence	Application
1	5'-TCCTCTCTTGGTTCTTCT-3'	EMSA
2	5'-TCTCTCG*CTCATCCAC-3'	EMSA
3	5'-TCTCTCFCTCATCCAC-3'	EMSA
4	5'-ACAGCGGYGCGAGGT-3'	EMSA
5	5'- GTCGTTCGGAAGACCCACTGCAAFGGGTTGATAATCGCCTCGCCATAG-3'	EMSA
6	5'-GCTCTACG*TCATCAC-3'	Crystallization/Crosslinking experiment
7	5'-TCTCTACGGTCATCAC-3'	Crystallization
8	5'-TCCTCTCTTG*TTCTTCT-3'	Crosslinking experiment
9	5'-CCGAGTCATTCTG* CAGCGAGTCCATGGGAGTC AAAT-3'	Crosslinking experiment
10	5'-GCTGACAGATCAGAFCAATGACTAGCTGAG-3'	Fluorescence polarization
11	5'-GCTCTACFTCATCAC-3'	Fluorescence polarization

F, FITC-dU; G*, AAF-dG; GG, 1,2-GG cisplatin; YY, CPD/ (6-4)PP.

	SeMet Rad14-AAF-dG	Rad14-AAF-dG	Rad14-1,2-GG
Wavelength, Å	0.979	0.976	0.873
Resolution range, Å	20–3.1 (3.2–3.1)	65.3–1.8 (1.84–1.80)	54.4–2.8 (2.9–2.8)
Space group	P 2 ₁	P 4 ₁	P 4 ₁
Unit cell	<i>a</i> = 66.4, <i>b</i> = 51.7, <i>c</i> = 68.8 β = 111.98	<i>a</i> = 53.8, <i>b</i> = 53.8 <i>c</i> = 130.4	<i>a</i> = 54.4, <i>b</i> = 54.4 <i>c</i> = 130.8
Total reflections	59,565 (5,963)	253,904 (14,638)	39,291 (3,836)
Unique reflections	7,284 (709)	34,232 (1948)	9,402 (933)
Multiplicity	8.2 (8.4)	7.4 (7.5)	4.2 (4.1)
Completeness, %	99.7 (99.2)	99.8 (95.6)	99.7 (98.1)
Mean <i>I</i> /σ(<i>I</i>)	21.3 (7.46)	10.7 (2.6)	9.1 (1.2)
Wilson B-factor	88.69	31.3	58.5
R-merge	0.07 (0.27)	0.08 (0.89)	0.1495 (1.33)
Anomalous completeness	0.995 (0.996)	98.8 (96)0.4	—
Anomalous multiplicity	4.1 (4.1)	3.4 (3.5)	—
CC _{1/2}	0.998 (0.994)	0.998 (0.593)	0.992 (0.485)
CC*	1 (0.999)	—	0.998 (0.808)
Figure of merit	0.33	—	—
R-work		0.21 (0.312)	0.213 (0.363)
R-free		0.241 (0.347)	0.256 (0.343)
Number of nonhydrogen atoms		3,172	4,088
Macromolecules		3,049	4,074
Ligands		80	14
Water		43	0
Protein residues		291	339
RMS (bonds)		0.018	0.012
RMS (angles)		2.15	1.50
Ramachandran favored, %		96	97
Ramachandran outliers, %		0	0
Average B-factor		48.6	55.7
Macromolecules		48.5	55.8
Ligands		60.4	48.3
Solvent		37	

Statistics for the highest-resolution shell are shown in parentheses.

Table S3. Table of melting points and thermodynamic parameters of the AAF-dG lesion containing duplex used for crystallization

Substrate	T _m , °C	Δ <i>H</i> , kcal/mol	Δ <i>S</i> , cal/mol K	Δ <i>G</i> _{298K} , kcal/mol
dG:dC	60	-73 ± 0.5	274	8.7
AAF-dG:dC	48	-70 ± 0.5	274	11.7

3.2. Strukturelle Basis für *bulky* DNA-Adduktschadenserkennung durch das Nukleotidexzisionsreparaturprotein Rad14

N. Simon*, C. Ebert*, S. Schneider, *Chemistry* **2016**, 22, 10782–10785. „*Structural Basis for Bulky-Adduct DNA-Lesion Recognition by the Nucleotide Excision Repair Protein Rad14*“. (* geteilte Erstautorenschaft)

3.2.1. Lizenz

Kopie der Publikation mit Erlaubnis des Verlags; Copyright 2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

3.2.2. Einführung

Die erste entstandene Kristallstruktur von Rad14 mit C8-AAF-dG geschädigter DNA zeigte (Abschnitt 3.1.), dass die Bindung des Rad14 Proteins zur Ausbildung eines DNA-Knicks von ungefähr 70° führt. Zudem ist der Schaden in der DNA Duplex an der Knickstelle interkaliert und die gegenüberliegende Base wird herausgeklappt. Die dabei möglichen π -stacking Interaktionen könnten für die Stabilisierung des DNA-Knicks eine bedeutende Rolle spielen. In diesem Manuskript wurde der Rad14-Erkennungsmechanismus weiter untersucht, insbesondere die π -stacking Interaktionen. Hierfür wurden verschiedene Schadensanaloga von AAF synthetisiert, um die Größe der π -stacking Interaktionen zu variieren, sowie in die jeweiligen Oligonukleotide eingebaut und mit Rad14 kristallisiert. Die erhaltenen Kristallstrukturen bestätigen den indirekten Schadenserkennungsmechanismus von Rad14, indem es die DNA auf Biegsamkeit und Flexibilität prüft. Zudem zeigt die Publikation, dass ein größerer stacking-Bereich des C8-Schadens zu einer besseren Stabilisierung des DNA-Knicks führt.

3.2.3. Autorenbeitrag

Für dieses Manuskript entwickelte ich eine neue Synthese für die C8-bulky-Schäden (C8-AAB-dG, C8-AAN-dG, C8-AAA-dG und C8-AAP). Der Einbau dieser Schäden in Oligonukleotide mittels DNA-Festphasensynthese wurde in Zusammenarbeit mit *Charlotte Ebert* durchgeführt. Weiterhin reinigte ich das Rad14 Protein für die Kristallisationsstudien auf. Die Schmelzpunktmessungen und

Kristallisationsstudien wurden zusammen mit *Charlotte Ebert* durchgeführt. Sabine Schneider übernahm die Auswertung der Kristallstrukturdaten.

Protein Crystal Structures

Structural Basis for Bulky-Adduct DNA-Lesion Recognition by the Nucleotide Excision Repair Protein Rad14

Nina Simon^{+, [b]} Charlotte Ebert^{+, [b]} and Sabine Schneider^{*, [a]}

Abstract: Heterocyclic aromatic amines react with purine bases and result in bulky DNA adducts that cause mutations. Such structurally diverse lesions are substrates for the nucleotide excision repair (NER). It is thought that the NER machinery recognises and verifies distorted DNA conformations, also involving the xeroderma pigmentosum group A and C proteins (XPA, XPC) that act as a scaffold between the DNA substrate and several other NER proteins. Here we present the synthesis of DNA molecules containing the polycyclic, aromatic amine C8-guanine lesions acetylaminophenyl, acetylaminonaphthyl, acetylaminanthryl, and acetylaminopyrenyl, as well as their crystal structures in complex with the yeast XPA homologue Rad14. This work further substantiates the indirect lesion-detection mechanism employed by the NER system that recognises destabilised and deformable DNA structures.

Defects in the nucleotide excision repair are known to be associated with several human, autosomal, recessive hereditary disorders, such as xeroderma pigmentosum (XP) and Cockayne Syndrome.^[1] Moreover, single-nucleotide polymorphisms (SNPs) in nucleotide excision repair (NER) genes have been associated with the susceptibility of cancers for the treatment with chemotherapeutic agents that exert their anti-tumour properties by inflicting DNA lesions.^[2] More than 30 proteins are involved in NER, which operates by a sequential assembly process. This comprises the recognition of damaged sites, excision of the oligonucleotide containing the lesion, gap filling by templated DNA synthesis and final sealing of the nick. The NER system is extremely versatile regarding the chemical structures of DNA lesions by either scanning the genome (global genome (GG)-NER) or acting on them during transcription (transcription-coupled NER).^[3] Currently it is believed that lesions are first recognised by the XPC (xeroderma pigmentosum

group C) protein^[4] followed by lesion verification by XPD (xeroderma pigmentosum group D protein), TFIIH (transcription factor IIH/XPB),^[5] XPA (xeroderma pigmentosum group A protein), replication protein A (RPA)^[6] and XPG (xeroderma pigmentosum group G protein),^[3c,e,7] resulting in the fully assembled pre-incision complex.^[3c,e,7] XPC, XPA and RPA were shown to ensure the proper 3D arrangement of the DNA intermediate for incision of the damaged single strand during the GG-NER.^[6] The interaction between RPA and XPA are essential for the NER.^[8] In addition, it was shown that XPA is a crucial component for both GG-NER as well as TC-NER,^[9] and mutations in XPA are amongst the strongest NER-phenotypes.^[10] Thus, XPA most likely acts as a scaffold during the NER.^[3e] Moreover, it was previously reported that XPA binds to a variety of DNA structures, such as kinks,^[11] DNA duplexes containing bulky adducts^[12] and cisplatin lesions.^[13]

To investigate the structural basis for the binding of the XPA homologue Rad14 from *Saccharomyces cerevisiae* to different C8-acetylated deoxyguanosine (dG) bulky-adduct lesions, we synthesised the acetylaminophenyl (dG-AAB), acetylaminonaphthyl (dG-AAN), acetylaminanthryl (dG-AAA) and acetylaminopyrenyl (dG-AAP) phosphoramidites and incorporated the lesions into 15-mer and 34-mer DNA oligonucleotides by ultramild solid-phase DNA synthesis (Figure 1 a, for details see the Supporting Information). The synthesis of the phosphoramidites dG-AAB, dG-AAA and dG-AAP was optimised by using a 2-(trimethylsilyl)ethanol (TMS) protecting group at the O⁶-position. TMS moieties can be removed in one step together with the *tert*-butyldimethylsilyl-ether (TBS) protecting groups.

The DNA duplexes containing the dG-AAB, dG-AAN, dG-AAA, or dG-AAP lesions were co-crystallised with the Rad14 minimal DNA-binding domain (DBD, residues 188–306) and the crystal structures determined by X-ray diffraction. (For data collection and structure refinement statistics see Table S3 in the Supporting Information).

As previously observed for the Rad14 binding to DNA with C8-dG-acetylaminofluorenyl (dG-AAF) or cisplatin (1,2 GG) lesions,^[14] two Rad14 protein molecules bind one DNA double strand with the DNA bent by about 75° into the major groove. The two proteins only interact with the DNA phosphodiester backbone and bury a solvent-accessible surface area of about 700 Å² per Rad14 molecule, but do not form any DNA-sequence- or DNA-lesion-specific contacts (Figure 1). In all complexes of Rad14 with damaged DNA, the bending of the DNA is achieved through the “fingers” with Gln266, Lys229 and Thr230, as well as His258 and Phe262 located on the β-hairpin, providing the anchor point by packing on the 13-mer recogni-

[a] Dr. S. Schneider

Center for Integrated Protein Science Munich CIPSM
Department of Chemistry, Technische Universität München
Lichtenbergstrasse 4, 85748 Garching (Germany)
E-mail: sabine.schneider@mytum.de[b] N. Simon,⁺ C. Ebert⁺Center for Integrated Protein Science Munich CIPSM
Department of Chemistry, Ludwig-Maximilians Universität München
Butenandtstrasse 13, 81377 München (Germany)

[†] These authors contributed equally to this work.

Supporting information for this article is available on the WWW under
<http://dx.doi.org/10.1002/chem.201602438>.

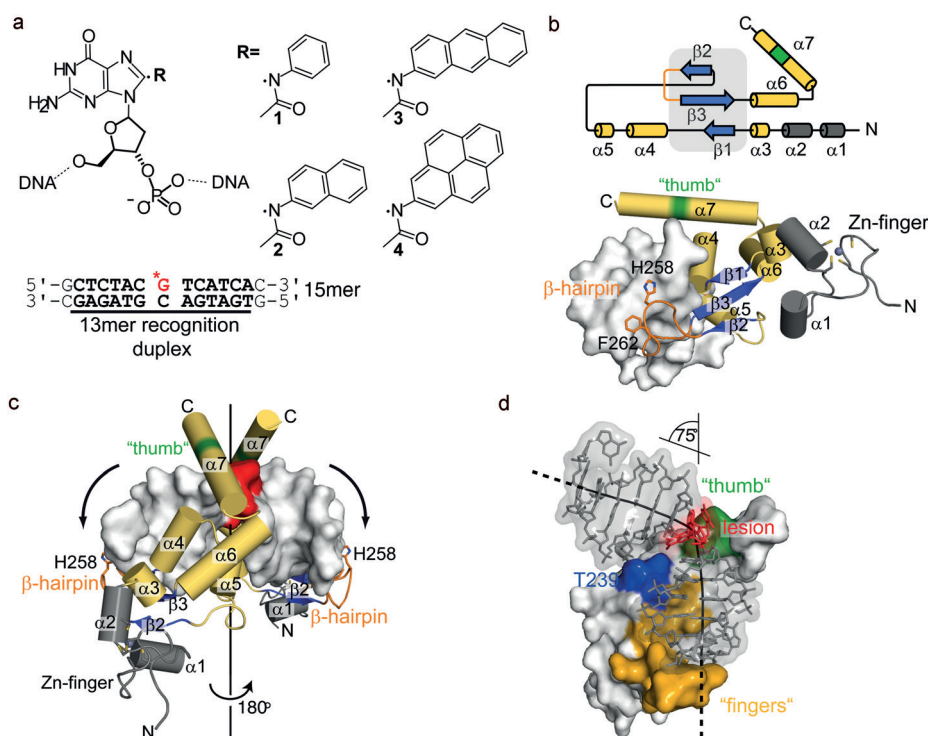


Figure 1. Structure of the Rad14 DBD in complex with DNA-containing bulky adducts. a) C8-dG-adduct lesions and DNA sequences used in this study. R = acetylaminophenyl (AAB, 1), acetylaminonaphthyl (AAN, 2), acetylaminanthryl (AAA, 3), acetylaminopyrenyl (AAP, 4). b) Overall folding topology of Rad14 and interaction with the DNA (grey surface). For clarity, only one Rad14 protein is shown. c) The DNA is bound by two Rad14 proteins and kinked at the position of the DNA lesion (red) towards the major groove (arrows). The protein is shown as cartoon and the DNA represented as grey surface. The β -hairpin with His258 and Phe262 (orange) pack on the 13-mer recognition duplex (the arginine "thumb" at the α 7 helix is highlighted in green). His258 and Phe262, as well as the Cys residues of the Zn-finger (dark grey), are shown as sticks and the Zn^{2+} as grey sphere. The black line indicates the pseudo C_2 symmetry axis. d) Rad14-DNA interface and bending mechanism; the "fingers" (gold) interact with the DNA backbone 3-6 base pairs away from the DNA-damage site. The "thumb" (green) bends the DNA at the lesion site by about 75° and Thr239 (blue) stabilises the kink. For clarity, only one Rad14 protein is depicted.

tion duplex. The arginine "thumb" (Arg293; Arg294) in the α 7 helix bends the DNA duplex with Thr239 stabilising the kinked structure (Figure 1). For the human XPA, it could recently be shown that the DNA-binding affinity was increased when the corresponding DBD (residues 98-219) was extended at the C terminus by 20 amino acids.^[15] This addition would lead to an extension at the α 7 helix resulting in a larger interaction area with the DNA duplex, which explains the higher binding affinity. All lesions are located at the central kink and the six base pairs on both sides adopt a B-form DNA topology that forms the 13-mer duplex-recognition element, in which the following bases of the DNA strand are unpaired (Figure 1). All structures exhibit a pseudo C_2 symmetry, in which the DNA is bound in two orientations, which is reflected by the mixed electron density observed for the purine-pyrimidine bases (Figure S2a in the Supporting Information). Thus, both orientations of the DNA duplexes were modelled with an occupancy of 50%. The dC (deoxycytidine) opposite of the damaged base is flipped out and disordered, while the lesions slide into the duplex and stabilise the kinked structure. The damaged guanine base is in *syn* conformation and displaced into the narrowed major groove. In the structure of Rad14 in complex with dG-AAP, the planes of the pyrene moieties in the two duplex orientations are almost perpendicular to each other, stacking planar with its respect 5'-flanking basepair and verti-

cal on the 3'-flanking base pair (Figure 2 and Figure S2 in the Supporting Information). The phenyl, naphthyl and anthryl moieties in the other complex structures are also fully intercalated into the duplex, albeit their positions are less well defined in the electron-density maps (Figure S1 in the Supporting Information), which is most likely due to the following inherent structural properties: 1) the two alternative DNA orientations result in only 50% occupancy for each lesion-containing strand and 2) the C_2 pseudo symmetry axis runs directly through the centre of the dsDNA at the position of the kink (Figure 1c) and causes noise in the electron-density maps, which is commonly observed along symmetry axis in protein crystals. Moreover,

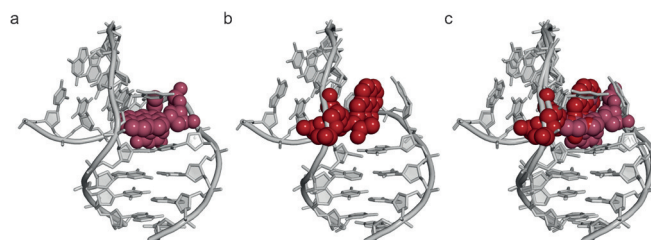


Figure 2. a, b) Position and stacking interactions of the dG-AAP lesions in the two orientations of the DNA duplex. c) Superposition of both orientations each with 50% occupancy. The dG-AAP lesions are shown as space-fill model shaded in grey.

the smaller aromatic moieties allow a higher degree of flexibility compared to the pyrene ring system, which will provide a better stabilisation of the DNA kink due to its larger stacking area. This is also reflected by the higher temperature factors of the lesions and their neighbouring base pairs (1.2–1.5 times higher compared to the average B-factors of the DNA strands).

Previously, it was shown that DNA lesions, which do not disturb the DNA-duplex structure or neighbouring Watson–Crick base pairs and/or increase the DNA stability (such as bulky *cis*- and *trans*-benzo[*a*]pyrene-*N*²-dG adducts placed opposite an abasic site), evade the NER mechanism and, in particular, the recognition by XPC.^[16] In general, good substrates for the NER system are DNA strands with disturbed Watson–Crick base pairs, kinked/bend and destabilised DNA-duplex structures,^[17] some of which have been shown to be bound by XPC and XPA.^[11, 12b, 13]

Therefore, the mechanism underlying the structural variety recognised by the NER can be explained if the system employs a sequence-independent shape-readout mechanism in its lesion recognition and verification step. An overall shape-readout mechanism is also used by a number of other DNA-binding proteins, such as Holliday junction-resolving enzymes, Z-DNA-recognising and bacterial chromosomal proteins. Any complex that forms favourable enthalpy changes needs to drive unfavourable entropy changes and vice versa.^[18] The lesions used in this study destabilise DNA and reduce the melting temperature compared to the undamaged control strand, albeit the overall destabilising effect is compensated in longer DNA duplexes (66–76% 15-mer; 91–98% 34-mer; Table S2). Due to the destabilisation effect of the DNA lesions, the enthalpic cost for the DNA bending is reduced, hence, a productive recognition complex can be formed. Analogously, in the mismatch repair, the initial mispair-recognition steps are binding and bending of the linear mispair-containing DNA by roughly 50–60°.^[19] In case of *N*² adducts placed opposite an abasic site that lack pre-deformation of the DNA and possess an increased stability due to full intercalation of the aromatic moiety, the enthalpic cost that is associated with binding by NER proteins is likely to high to be compensated by favourable enthalpic and entropic changes. Hence, these lesions are resistant to the NER. Such an indirect shape-readout mechanism is also supported by the observations made by Reardon and Sancar; RPA, XPA and XPC recognise psoralen damage-containing DNA, but only RPA can be crosslinked to the lesion.^[20] The requirement for this chemical crosslink is the close proximity (ca. 8 Å) of the photoreactive amino acids to the pyrene side ring.^[21] Employing a nonspecific shape-readout mechanism, the lesion recognition and verification is likely to enable the NER repair machinery to deal with this vast variety of different chemical structures. Our structural data on the Rad14 DBD in complex with these bulky aromatic dG-adduct lesions further circumstantiates the hypothesis that Rad14, and most likely its human homologue XPA (sharing a 25% sequence identity and a 45% sequence similarity), is able to recognise numerous different lesions in the lesion-verification step of the NER by utilising a general mechanism of binding bendable DNA or DNA adopting already bend conformations.

Experimental Section

Phosphoramidite synthesis and DNA preparation

The dG-AAN phosphoramidite containing an isopropylphenoxyacetyl group at the *N*²-position, for the solid-phase DNA synthesis, was synthesised as previously published.^[22] For the synthesis of the dG-AAB-dG, dG-AAA and dG-AAP phosphoramidites, instead of the benzyl ether a 2-(trimethylsilyl)-ethanol (TMS) moiety was used at the *O*⁶-position, which could be removed in one step with the *tert*-butyldimethylsilyl groups, enabling a faster synthesis. The C8-lesion phosphoramidites were incorporated into DNA by solid-phase DNA synthesis under ultra-mild conditions.^[23] For synthetic and purification details, and a list of oligodesoxynucleotides (ODN) used in this study, see the Supporting Information.

Protein expression, purification, crystallisation and structure determination

The *Saccharomyces cerevisiae* Rad14 DNA-binding domain (residues 188–306) was expressed and purified as published elsewhere.^[14] For details on the complex crystallisation and structure determination see the Supporting Information. Atomic coordinates and X-ray structure data were submitted to the Protein Data Bank (<http://www.ebi.ac.uk/pdbe/>). PDB codes: 5G32, 5G33, 5G34 and 5G35.

Acknowledgements

The authors would like to thank the crystallisation facility at the Max Planck Institute of Biochemistry for setting up the crystallisation plates. We also want to thank the Swiss Light Source (SLS) and European Synchrotron radiation Facility (ESRF) for beamtime and the staff of beamlines PXI (SLS) and ID23-2, ID29, and ID30B (ESRF) for assistance with the crystal testing and data collection. We would also like to thank Thomas Carell for helpful discussions and critical reading of the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft, the excellence cluster CIPSM and the Fonds der Chemischen Industrie.

Keywords: bulky adducts • crystal structures • DNA damage • nucleotide excision repair • Rad14

- [1] a) M. Berneburg, J. Krutmann, *Hautarzt* **2003**, *54*, 33–40; b) J. E. Cleaver, *J. Dermatol. Sci.* **2000**, *23*, 1–11; c) C. F. Menck, V. Munford, *Genet. Mol. Biol.* **2014**, *37*, 220–233.
- [2] M. Monzo, I. Moreno, A. Navarro, R. Ibeas, R. Artells, B. Gel, F. Martinez, J. Moreno, R. Hernandez, M. Navarro-Vigo, *Oncology* **2008**, *72*, 364–370.
- [3] a) S. Alekseev, F. Coin, *Cell. Mol. Life Sci.* **2015**, *72*, 2177–2186; b) E. C. Friedberg, *Nat. Rev. Cancer* **2001**, *1*, 22–33; c) J. A. Martijn, H. Lans, W. Vermeulen, J. H. Hoeijmakers, *Nat. Rev. Mol. Cell Biol.* **2014**, *15*, 465–481; d) I. Mellon, G. Spivak, P. C. Hanawalt, *Cell* **1987**, *51*, 241–249; e) H. Naegeli, K. Sugawara, *DNA Repair* **2011**, *10*, 673–683; f) A. Sancar, *Science* **1994**, *266*, 1954–1956; g) S. Tornaletti, P. C. Hanawalt, *Biochimie* **1999**, *81*, 139–146.
- [4] a) M. S. Luijsterburg, G. von Bornstaedt, A. M. Gourdin, A. Z. Politi, M. J. Mone, D. O. Warmerdam, J. Goedhart, W. Vermeulen, R. van Driel, T. Hofer, *J. Cell Biol.* **2010**, *189*, 445–463; b) J. H. Min, N. P. Pavletich, *Nature* **2007**, *449*, 570–575; c) T. Riedl, F. Hanaoka, J. M. Egly, *EMBO J.* **2003**, *22*, 5293–5303; d) M. Volker, M. J. Mone, P. Karmakar, A. van Hofen, W. Schul, W. Vermeulen, J. H. Hoeijmakers, R. van Driel, A. A. van Zeeland, L. H. Mullenders, *Mol. Cell* **2001**, *8*, 213–224.

- [5] a) J. Kuper, S. C. Wolski, G. Michels, C. Kisker, *EMBO J.* **2012**, *31*, 494–502; b) N. Mathieu, N. Kaczmarek, H. Naegeli, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 17545–17550; c) N. Mathieu, N. Kaczmarek, P. Ruthemann, A. Luch, H. Naegeli, *Curr Biol* **2013**, *23*, 204–212; d) R. A. Pugh, C. G. Wu, M. Spies, *EMBO J.* **2012**, *31*, 503–514.
- [6] Y. S. Krasikova, N. I. Rechkunova, E. A. Maltseva, I. O. Petruseva, O. I. Lavrik, *Nucleic Acids Res.* **2010**, *38*, 8083–8094.
- [7] a) K. Sugawara, J. Akagi, R. Nishi, S. Iwai, F. Hanaoka, *Mol. Cell* **2009**, *36*, 642–653; b) A. Tapias, J. Auriol, D. Forget, J. H. Enzlin, O. D. Scharer, F. Coin, B. Coulombe, J. M. Egly, *J. Biol. Chem.* **2004**, *279*, 19074–19083.
- [8] L. Li, X. Lu, C. A. Peterson, R. J. Legerski, *Mol. Cell. Biol.* **1995**, *15*, 5396–5402.
- [9] T. Kobayashi, S. Takeuchi, M. Saijo, Y. Nakatsu, H. Morioka, E. Otsuka, M. Wakasugi, O. Nikaido, K. Tanaka, *Nucleic Acids Res.* **1998**, *26*, 4662–4668.
- [10] I. Satokata, K. Tanaka, S. Yuba, Y. Okada, *Mutat. Res.* **1992**, *273*, 203–212.
- [11] U. Camenisch, R. Dip, S. B. Schumacher, B. Schuler, H. Naegeli, *Nat. Struct. Mol. Biol.* **2006**, *13*, 278–284.
- [12] a) N. E. Geacintov, S. Broyde, T. Buterin, H. Naegeli, M. Wu, S. Yan, D. J. Patel, *Biopolymers* **2002**, *65*, 202–210; b) Y. Liu, Z. Yang, C. Utzat, G. Wang, A. K. Basu, Y. Zou, *Biochemistry* **2005**, *44*, 7361–7368; c) Z. G. Yang, Y. Liu, L. Y. Mao, J. T. Zhang, Y. Zou, *Biochemistry* **2002**, *41*, 13012–13020.
- [13] a) E. Evans, J. Fellows, A. Coffey, R. D. Wood, *EMBO J.* **1997**, *16*, 625–638; b) C. J. Jones, R. D. Wood, *Biochemistry* **1993**, *32*, 12096–12104; c) M. Missura, T. Buterin, R. Hindges, U. Hubscher, J. Kasparkova, V. Brabec, H. Naegeli, *EMBO J.* **2001**, *20*, 3554–3564.
- [14] S. C. Koch, J. Kuper, K. L. Gasteiger, N. Simon, R. Strasser, D. Eisen, S. Geiger, S. Schneider, C. Kisker, T. Carell, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 8272–8277.
- [15] N. Sugitani, S. M. Shell, S. E. Soss, W. J. Chazin, *J. Am. Chem. Soc.* **2014**, *136*, 10830–10833.
- [16] a) Z. Liu, S. Ding, K. Kropachev, L. Jia, S. Amin, S. Broyde, N. E. Geacintov, *PLoS One* **2015**, *10*, e0137124; b) D. A. Reeves, H. Mu, K. Kropachev, Y. Cai, S. Ding, A. Kolbanovskiy, M. Kolbanovskiy, Y. Chen, J. Krzeminski, S. Amin, *Nucleic Acids Res.* **2011**, *39*, 8752–8764.
- [17] a) Y. Cai, D. J. Patel, N. E. Geacintov, S. Broyde, *J. Mol. Biol.* **2007**, *374*, 292–305; b) K. Kropachev, M. Kolbanovskii, Y. Cai, F. Rodriguez, A. Kolbanovskii, Y. Liu, L. Zhang, S. Amin, D. Patel, S. Broyde, *J. Mol. Biol.* **2009**, *386*, 1193–1203.
- [18] a) S. Harteis, S. Schneider, *Int. J. Mol. Sci.* **2014**, *15*, 12335–12363; b) L. Jen-Jacobson, L. E. Engler, L. A. Jacobson, *Structure* **2000**, *8*, 1015–1023; c) R. Rohs, X. Jin, S. M. West, R. Joshi, B. Honig, R. S. Mann, *Annu. Rev. Biochem.* **2010**, *79*, 233–269; d) A. A. Travers, C. Vaillant, A. Arneodo, G. Muskhelishvili, *Biochem. Soc. Trans.* **2012**, *40*, 335–340.
- [19] G. L. Hura, C. L. Tsai, S. A. Claridge, M. L. Mendillo, J. M. Smith, G. J. Williams, A. J. Mastroianni, A. P. Alivisatos, C. D. Putnam, R. D. Kolodner, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 17308–17313.
- [20] J. T. Reardon, A. Sancar, *Mol. Cell. Biol.* **2002**, *22*, 5938–5945.
- [21] S. S. Sastry, *Photochem. Photobiol.* **1997**, *65*, 937–944.
- [22] a) Z. Wang, C. J. Rizzo, *Org. Lett.* **2001**, *3*, 565–568; b) L. C. Gillet, O. D. Scharer, *Org. Lett.* **2002**, *4*, 4205–4208.
- [23] L. C. Gillet, J. Alzeer, O. D. Scharer, *Nucleic Acids Res.* **2005**, *33*, 1961–1969.

Received: May 23, 2016

Published online on June 27, 2016

CHEMISTRY

A **European** Journal

Supporting Information

Structural Basis for Bulky-Adduct DNA-Lesion Recognition by the Nucleotide Excision Repair Protein Rad14

Nina Simon^{+, [b]} Charlotte Ebert^{+, [b]} and Sabine Schneider^{*[a]}

chem_201602438_sm_miscellaneous_information.pdf

Supporting Information

Structural basis for bulky adduct DNA lesion recognition by the nucleotide excision repair protein Rad14

Nina Simon, Charlotte Ebert and Sabine Schneider

General Methods

Chemicals were purchased from Sigma-Aldrich, Fluka, Alfa Aesar, ChemGenes or Acros and used without further purification. Solutions were concentrated in vacuo on a Heidolph rotary evaporator. The solvents were of reagent grade and purified by distillation. Dry solvents were bought from Sigma-Aldrich. Acetonitrile for HPLC-purification was purchased from VWR. Water was purified by a Milli-Q Plus system from Merck Millipore. Chromatographic purification of products was accomplished using flash column chromatography on Merck Geduran Si 60 (40-63 μ M) silica gel (normal phase). Thin layer chromatography (TLC) was performed on Merck 60 (silica gel F₂₅₄) plates. ¹H, ¹³C and ³¹P NMR spectra were recorded in deuterated solvents on *Bruker ARX 300*, *Varian VXR400S*, *Varian Inova 400*, *Bruker AMX 600* and *Bruker Avance 800* spectrometers and calibrated to the residual solvent peak. Multiplicities are abbreviated as follows: s = singlet, d = doublet, t = triplet, q = quartet, hept = septet, m = multiplet, br. = broad. For assignment of the structures, additional 2D NMR spectra (COSY, HSQC, HMBC) were measured. Matrix-assisted laser desorption/ionization-time-of-flight (MALDI-TOF) mass spectra were recorded on a Bruker Autoflex II. For MALDI-TOF measurements the samples were dialysed on a 0.025 μ m VSWP filter (Millipore) against ddH₂O for 1 h and then cocrystallised in a 3-hydroxypicolinic acid matrix (HPA: 25 mg 3-hydroxypicolinic acid, 5 mg ammonium citrate, 5 μ L 15-crown-5 in 0.5 mL H₂O). High resolution electrospray ionization mass spectra (HRMS-ESI) were recorded on a Thermo Finnigan LTQ-FT (ESI-FTICR). The pH-values of buffers were adjusted using a MP 220 pH-meter (Mettler Toledo). UV spectra and melting profiles were measured on a Jasco V-650 spectrometer using quartz glass cuvettes with 1 cm path length. The concentration of the purified DNA was determined with a NanoDrop ND-1000 spectrophotometer (Pqlab) from Thermo Scientific (1.5 μ L sample volume) using calculated extinction coefficients from Oligo Calc (online: <http://www.basic.northwestern.edu/biotools/OligoCalc.html>) and applying Lambert-Beer's law. For strands containing artificial bases the extinction coefficient of their corresponding control strand was employed without corrections.

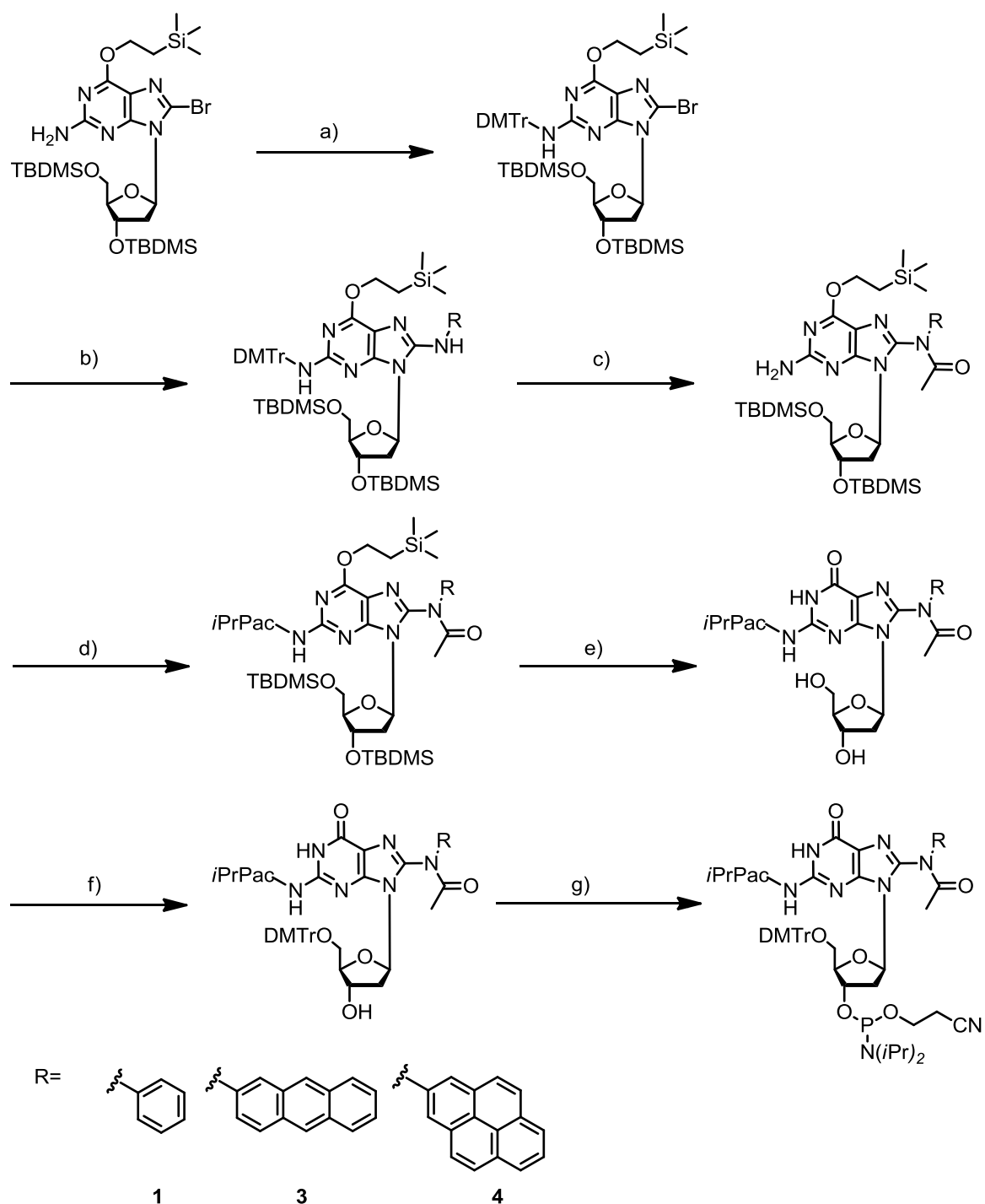
Phosphoramidite and DNA synthesis Phosphoramidites for dA, dC, dG, dT and CPG carriers were obtained from Glen Research (Sterling, VA) or Link Technologies (Bellshill, Scotland). Incorporation of the C8-lesion phosphoramidites into DNA was done by solid phase DNA synthesis with ultra-mild conditions,^[1] using an ABI 394 Nucleic Acid Synthesis System (Life Technologies, Darmstadt, Germany). The coupling time for the modified phosphoramidites was extended to 3 x 15 min. The oligonucleotides were purified by reverse phase HPL chromatography (Macherey-Nagel, Nucleosil 100-7 C18, 10 x 250 mm, 0.1 M triethylammonium acetate, linear gradient from 0% to 40% in 45 min for the 15mers or 0% to 20% acetonitrile in 65 min for the longer oligonucleotides). The purified oligonucleotides were desalted using C18 cartridges (Waters, Sep-Pak C18 Classic Cartridge) and their identity confirmed by matrix-assisted laser-desorption time-of-flight mass spectrometry.

Determination of melting temperatures The DNA duplexes (2 μ M) were dissolved in 250 μ L of 10 mM Tris-HCl, pH 7.4, 150 mM NaCl. The ODN 1 and ODN 3 samples were heated from 15 to 90 °C or cooled from 90 to 15 °C with a rate of 1 °C/min. The ODN 2 and ODN 4 samples were heated from 20 to 95 °C or cooled from 95 to 20 °C with a rate of 1 °C/min. Three melting profiles per sample were performed.

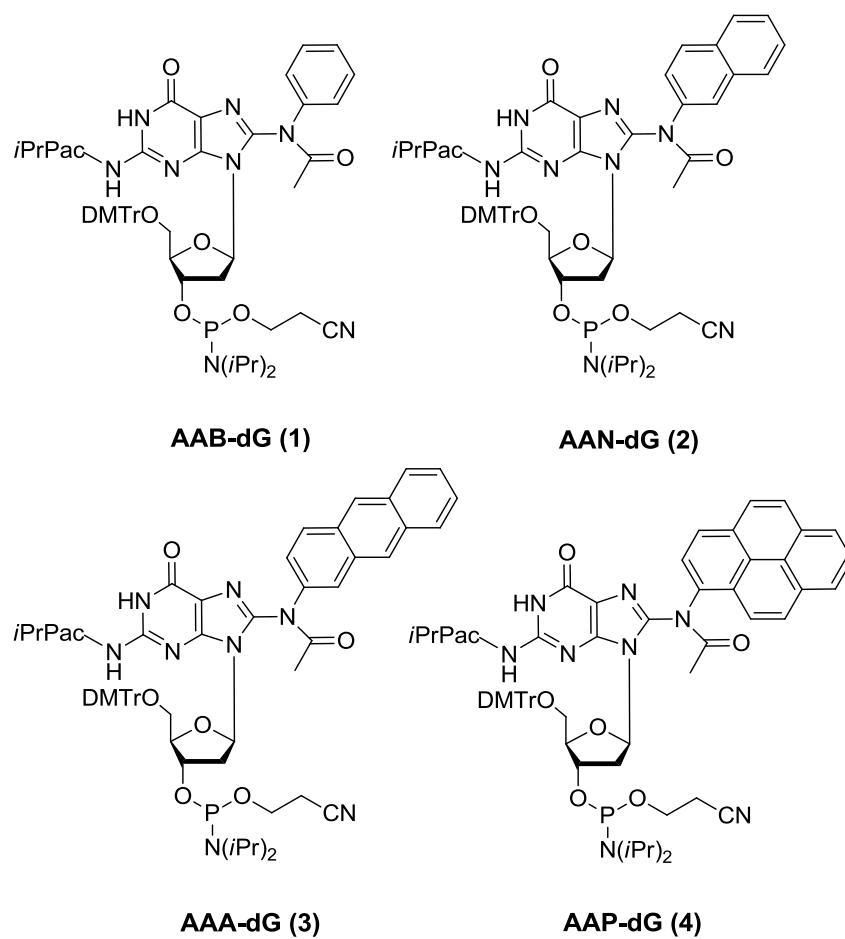
Data processing and structure determination

The *Saccharomyces cerevisiae* Rad14 DNA binding domain (residues 188-306) was expressed and purified as published in.^[2] The C8-lesion-dG containing 15mer DNA (ODN1, see Supplementary Table 1) was annealed in crystallisation buffer to its counter strand (ODN 2, Metabion). Prior to crystallisation protein and DNA were mixed in a 2:1 (protein:DNA) molar ratio and incubated for 30 min at 4 °C. Crystals were grown in 0.2-0.22 M ammonium nitrate and 38 - 44% 2-methyl-1,3,-propanediol (MPD) at 4 °C, cryo-cooled in liquid nitrogen directly from the crystallisation drop and stored in liquid nitrogen until data collection. Due to the small crystal size of 10-20 μ m, diffraction data were collected with a beam size of 10 x 10 μ m (PX I beamline, Swiss Light Source (SLS); ID30B beamline, European Synchrotron Radiation Facility (ESRF)). The data were processed with XDS^[3] and the resolution cut-offs were chosen according to the correlation coefficient of random half-data sets (1/2 CC) at about 50%.^[4] The structures were solved using the coordinates of the Rad14-DNA complex (PDB code 5A3D) by rigid-body refinement in REFMAC5^[5] after the atoms for the DNA were removed from the model. The models were completed by iterative cycles of manual model building in COOT^[6] and restraint and TLS refinement^[7] in REFMAC5. The pseudo C2 two fold symmetry axis that runs through the centre of the DNA double strand can be turned into a crystallographic two-fold axis by processing the data with comparable statistics in the space group P4₁2₁2, instead of P4₁. However this results in a composition of the crystals asymmetric unit (asu) of one protein with one dsDNA strand at 50% occupancy (P4₁2₁2)

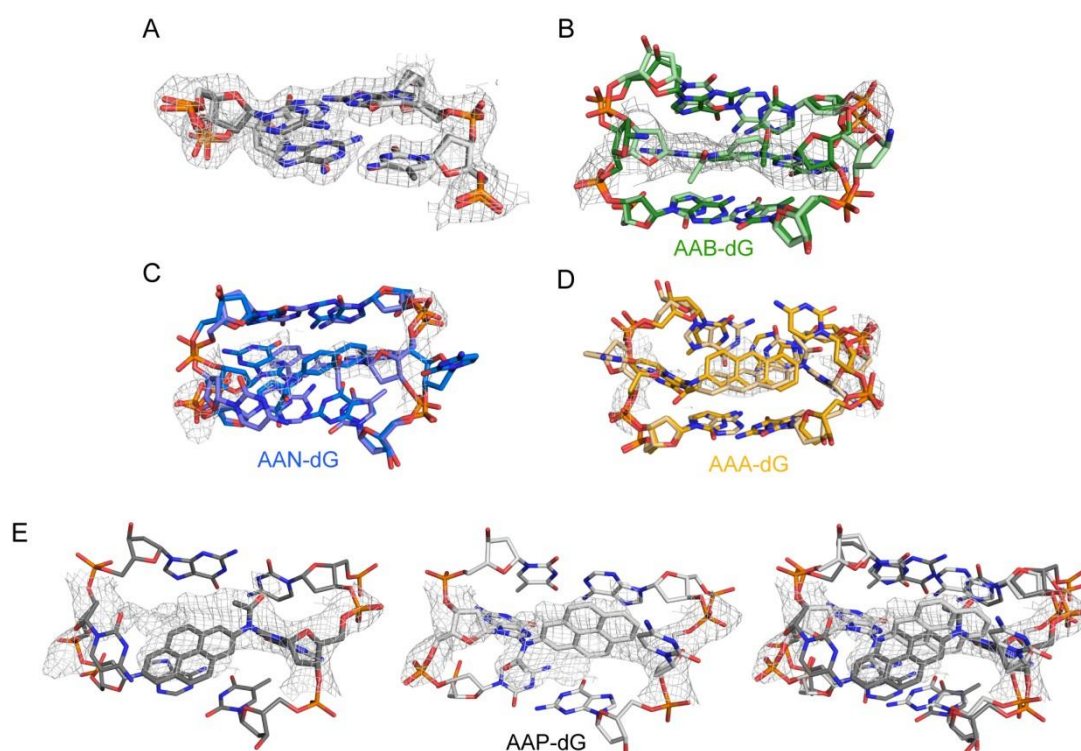
compared to two proteins with two dsDNA each with 50% occupancy (P4₁). For both space groups comparable model refinement statistics were obtained. Thus the crystal system P4₁ was chosen, which contains the biological relevant complex of two proteins recognising one dsDNA in two orientations in the asu. Simulated-annealing omit electron density map coefficients were calculated using PHENIX.^[8] Structure validation and analysis was carried out with MOLPROBITY^[9], PISA^[10] and CURVES.^[11] For data processing and structure refinement details and statistics see the Supporting Information. Structural superposition were done with SSM.^[12] All structural figures were prepared with PyMol (Delano Scientific).



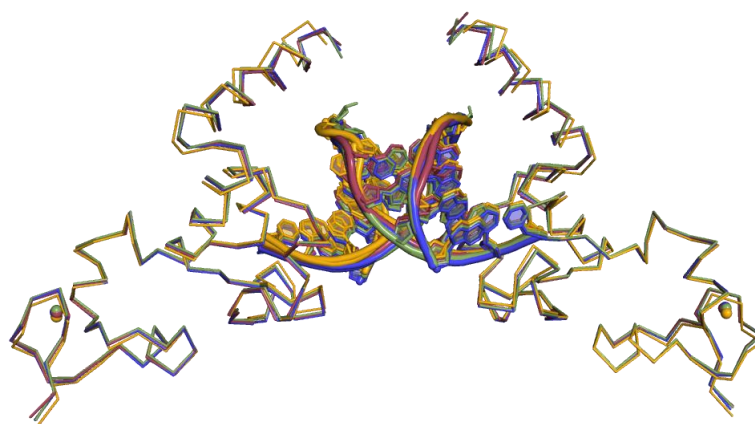
Supporting Scheme 1. Synthesis of the phosphoramidites **1**, **3** and **4**. a) DMTr-Cl, pyridine, RT, b) Pd₂db₃, BINAP, NaOtBu, toluol, aryl-NH₂ (1,2 or 3), 110 °C, c) 1. 4-DMAP, NEt₃, Ac₂O, pyridine, RT, 2. HCl in MeOH (0.01 M), RT, d) *i*PrPac-Cl, pyridine, RT, e) TBAF, AcOH, THF, RT, f) DMTr-Cl, pyridine, RT, g) (*i*Pr₂N)₂PO(CH₂)₂CN, DIAT, DCM, RT.



Supporting Figure 1: Synthesized C8-dG lesion phosphoramidites.



Supporting Figure 2. Illustration of the electron density maps. A) Observed 2Fo-DFc electron density contoured at 1 σ , indicating the two DNA orientations. Simulated annealing Fo-DFc electron density maps, contoured at 2.5 σ , calculated excluding the atoms for dG-AAB (B), AAN (C), dG-AAA (D) and dG-AAP (E).



Supporting Figure 3. Structural superposition of all Rad14-DNA structures. The structures can be superimposed with an r.m.s.d of 0.3-0.8 Å depending on the chosen reference frame. (only main chain atoms of the protein vs all atoms) The dG-AAB complex is shown in green, the dG-AAN complex in blue, the dG-AAA complex in golden and the dG-AAP complex in red. The protein is depicted as ribbon, with the DNA shown as cartoon and the Zn²⁺ as sphere.

Supporting Table 1. Oligonucleotide sequences used in this study. The damaged guanine base is marked with a star and highlighted in bold.

ODN	Oligonucleotide sequence
1	5'-GCT CTA CG *T CAT CAC-3'
2	5'-GTA GCT AGG TCG AT G * CAC AGA CAT CGC ACA CGT C-3'
3	5'-GTG ATG ACG TAG AGC-3'
4	5'-GACG TGT GCG ATG TCT GTG CAT CGA CCT AGC TAC-3'

Supporting Table 2. Destabilising effect of the different lesions on the thermal stability of duplex DNA shown by the respective melting points.

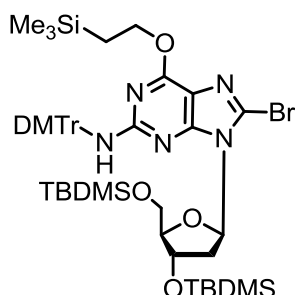
DNA lesion	T _M [°C] 15mer	T _M [°C] 34mer
dG:dC	64	84
C8-AAB-dG:dC	42	77
C8-AAN-dG:dC	46	77
C8-AAA-dG:dC	49	80
C8-AAP-dG:dC	47	80
C8-AAF-dG:dC	48	82

Supporting Table 3. Data processing and structure refinement statistics. Statistics for the highest-resolution shell are shown in parentheses.

	Rad14-dG-AAB	Rad14-dG-AAN	Rad14-dG-AAA	Rad14-dG-AAP
Wavelength	0.979	0.979	0.979	0.979
Resolution range	49.1 - 2.2 (2.3 - 2.2)	41.4 - 2.4 (2.5 - 2.4)	49.2 - 1.9 (2.0 - 1.9)	37.6 - 2.0 (2.1 - 2.0)
Space group	P 4 ₁	P 4 ₁	P 4 ₁	P 4 ₁
Unit cell	53.0 53.0 130.9	53.4 53.4 131.3	53.1 53.1 130.3	53.2 53.2 131.3
Total reflections	123,844 (11,546)	50,634 (4,920)	125,416 (8,792)	82,958 (8,405)
Unique reflections	18,268 (1,818)	14,071 (1,392)	28,204 (2,749)	24,353 (2,436)
Multiplicity	6.8 (6.4)	3.6 (3.5)	4.4 (3.2)	3.4 (3.5)
Completeness (%)	1.00 (1.00)	0.98 (0.98)	1.00 (0.97)	0.99 (0.99)
Mean I/sigma(I)	6.62 (0.97)	5.8 (1.1)	9.8 (1.4)	7.2 (1.0)
Wilson B-factor	41.30	46.9	33.0	36.0
R-merge	0.163 (1.111)	0.189 (0.994)	0.071 (0.681)	0.097 (1.248)
R-meas	0.177 (1.21)	0.223 (1.17)	0.081 (0.809)	0.115 (1.476)
CC1/2	0.995 (0.568)	0.976 (0.531)	0.998 (0.335)	0.997 (0.456)
CC*	0.999 (0.851)	0.994 (0.833)	1 (0.709)	0.999 (0.791)
Reflections used in refinement	18,255 (1,816)	14,056 (1,390)	28,199 (2,749)	24,338 (2,435)
Reflections used for R-free	895 (96)	720 (89)	1,361 (129)	1,255 (121)
R-work	0.215 (0.323)	0.255 (0.374)	0.214 (0.328)	0.214 (0.355)
R-free	0.253 (0.296)	0.307 (0.402)	0.245 (0.342)	0.2587 (0.392)
CC(work)	0.939 (0.700)	0.937 (0.739)	0.934 (0.613)	0.955 (0.608)
CC(free)	0.940 (0.793)	0.919 (0.523)	0.933 (0.464)	0.945 (0.678)
Number of non-hydrogen atoms	3,053	3,069	3,121	3,134
macromolecules	3,015	3,061	3,062	3,079
ligands	2	2	2	2
Protein residues	229	229	228	229
RMS(bonds)	0.042	0.031	0.052	0.034
RMS(angles)	1.80	1.72	1.87	2.12
Ramachandran favored (%)	96	95	98	98
Ramachandran allowed (%)	3.6	5.3	1.8	2.2
Ramachandran outliers (%)	0	0	0	0
Rotamer outliers (%)	2	3.4	0.98	0.97
Clashscore	2.5	2.1	1.1	3.5
Average B-factor	54.5	57.8	51.2	47.9
macromolecules	54.6	57.9	51.4	48.1
ligands	35.5	50.8	30.4	33.6
solvent	38.8	32.6	39.0	39.9
Number of TLS groups	20	20	20	28

Synthesis of phosphoramidites

C8-Br-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-*N*²-(4,4'-dimethoxytrityl)-*O*⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine



C8-Br-*O*⁶-(2-(trimethylsilyl)-ethyl)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-2'-deoxyguanosine^[13] (1.50 g, 2.23 mmol, 1.0 eq.) was dissolved in dry pyridine, molecular sieves (4 Å) were added and the solution was stirred at room temperature for 3 h. DMT-chloride (1.04 g, 2.68 mmol, 1.2 eq.) and 0.37 mL dry trimethylamine (2.7 mmol, 1.2 eq.) were added and the mixture was stirred at room temperature for 72 hours. The residue was purified via column chromatography (*iso*hexane 30:1 EtOAc + 1% pyridine; SiO₂) and C8-Br-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-*N*²-(4,4'-dimethoxytrityl)-*O*⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine (2.03 g, 2.07 mmol, 93%) was obtained as a white foam.

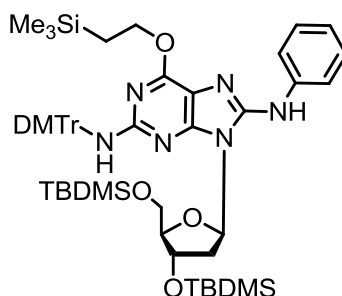
¹H-NMR (400 MHz, Acetone-*d*₆): δ = 7.36 (d, 3H, DMTr), 7.25 (d, 6H, DMTr), 6.84 (d, 4H, DMTr), 6.40 (s, *br*, 1H, NH), 6.23 (t, 1H, C1'-H), 4.84 (s, 1H, C3'-H), 3.98-3.81 (m, 4H, *O*⁶-CH₂, C5'-H, C4'-H), 3.77 (s, 6H, OCH₃), 3.68 (dd, 1H, C5'-H), 3.61-3.57 (m, 1H, C2'-H), 2.23-2.12 (m, 1H, C2'-H), 0.93 (s, 9H, C(CH₃)₃), 0.90–0.87 (m, 2H, *O*⁶-CH₂-CH₂), 0.84 (s, 9H, C(CH₃)₃), 0.16 (s, 3H, SiCH₃), 0.15 (s, 3H, SiCH₃), 0.00 (s, 3H, SiCH₃), -0.04 (s, 9H, Si(CH₃)₃), -0.05 (s, 3H, SiCH₃) *ppm*.

¹³C-NMR (101 MHz, Acetone-*d*₆): δ = 159.3, 158.5, 155.0, 147.2, 139.1, 131.0, 129.8, 129.0, 128.4, 127.4, 126.5, 116.9, 113.8, 113.7, 88.1, 86.7, 65.5, 63.5, 55.5, 36.6, 26.3, 26.2, 18.9, 18.7, 17.7, -1.4, -4.3, -4.4, -5.1, -5.1 *ppm*.

HR-MS (ESI⁺): *m/z* calculated for [C₄₈H₇₁BrN₅O₆Si₃]⁺ ([M+H]⁺): 976.3890, found: 976.3895.

The phosphoramidites of the pyrenyl and the anthryl adducts were prepared analogous to the phenyl adduct phosphoramidite.

8-(2-aminobenzene)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-*N*²-(4,4'-dimethoxytrityl)-O⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine



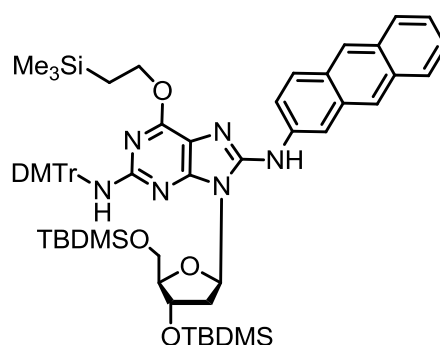
C8-Br-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-*N*²-(4,4'-dimethoxytrityl)-O⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine (2.00 g, 2.05 mmol, 1.0 eq.) was dissolved in 35 mL dry toluene under argon atmosphere. Aniline (0.37 mL, 4.1 mmol, 2.0 eq.), Tris(dibenzylidenacetone)dipalladium(0) (0.19 g, 0.21 mmol, 0.1 eq.) and (2,2'-bis(diphenylphosphino)-1,1'-binaphthyl) (0.39 g, 0.62 mmol, 0.3 eq.) were added and the reaction mixture was heated to 110 °C. Sodium-*tert*-butanolate (0.39 g, 4.1 mmol, 2.0 eq.) dissolved in 35 mL toluene was added and the reaction mixture was stirred at 110 °C for 1.5 h. After cooling down to room temperature the mixture was filtered over SiO₂ deactivated with pyridine and the filter cake was washed with 1% pyridine in EtOAc (150 mL). The filtrate was concentrated under reduced pressure and purified via column chromatography (isohexane/EtOAc 1:0 to 7:1 + 1% pyridine; SiO₂). 8-(2-aminobenzene)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-*N*²-(4,4'-dimethoxytrityl)-O⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine (1.62 g, 1.64 mmol, 80%) was obtained as a red brown foam.

¹H-NMR (400 MHz, Acetone-*d*₆): δ = 7.96 (s, *br*, 1H, C8-NH), 7.73–7.69 (m, 2H, C₆H₅), 7.41–7.37 (m, 3H, DMTr), 7.28 (d, 6H, DMTr), 7.25 (d, 2H, C₆H₅), 6.93 (tt, 1H, C₆H₅), 6.83 (d, 4H, DMTr), 6.23 (s, *br*, 1H, *N*²H), 6.16 (t, 1H, C1'-H), 4.70 (s, 1H, C3'-H), 4.07–3.98 (m, 2H, O⁶-CH₂), 3.97–3.92 (m, 2H, C5'-H, C4'-H), 3.82 (dd, 1H, C5'-H), 3.76 (s, 6H, OCH₃), 3.26–3.11 (m, 1H, C2'-H), 2.15 (ddd, 1H, C2'-H), 0.99–0.94 (m, 2H, O⁶-CH₂CH₂), 0.93 (s, 9H, C(CH₃)₃), 0.85 (s, 9H, C(CH₃)₃), 0.15 (s, 3H, SiCH₃), 0.14 (s, 3H, SiCH₃), 0.04 (s, 3H, SiCH₃), 0.01 (s, 3H, SiCH₃), 0.00 (s, 9H, Si(CH₃)₃) ppm.

¹³C-NMR (101 MHz, Acetone-*d*₆): δ = 159.1, 158.4, 157.0, 147.4, 141.6, 139.5, 131.1, 129.9, 129.7, 129.4, 128.3, 127.2, 124.6, 122.2, 119.0, 113.5, 113.3, 88.4, 85.0, 73.3, 70.9, 64.6, 63.8, 55.4, 38.7, 26.4, 26.3, 19.0, 18.7, 18.1, -1.3, -4.3, -4.4, -5.1, -5.1 ppm.

HR-MS (ESI⁺): *m/z* calculated for [C₅₄H₇₇N₆O₆Si₃]⁺ ([M+H]⁺): 989.5207, found: 989.5211.

8-(2-aminoanthracene)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-*N*²-(4,4'-dimethoxytrityl)-O⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine:



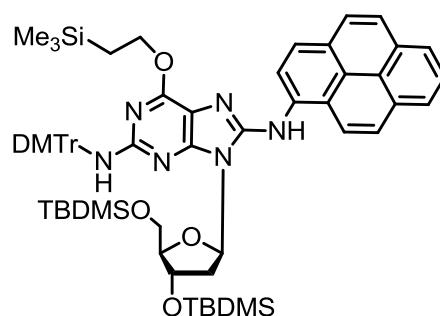
Yield: 64%

¹H-NMR (400 MHz, Acetone-*d*₆): δ = 8.42 (s, 1H, -C₁₄H₉), 8.32 (s, 1H, -C₁₄H₉), 8.28 (s, *br*, 1H, C8-NH), 7.99–7.96 (m, 3H, -C₁₄H₉), 7.79–7.69 (m, 1H, -C₁₄H₉), 7.63 (dd, 1H, -C₁₄H₉), 7.45–7.37 (m, 2H, -C₁₄H₉, 3H, DMTr), 7.30 (d, 6H, DMTr), 6.84 (d, 4H, DMTr), 6.28 (s, 1H, *N*²H), 6.23 (t, 1H, C1'-H), 4.77–4.71 (m, 1H, C3'-H), 4.08–4.05 (m, 2H, O⁶-CH₂), 4.02–3.96 (m, 2H, C5'-H, C4'-H), 3.86 (dd, 1H, C5'-H), 3.77 (s, 6H, OCH₃), 3.24 (dd, 1H, C2'-H), 2.21 (dt, 1H, C2'-H), 1.00 (dd, 2H, O⁶-CH₂CH₂), 0.93 (s, 9H, C(CH₃)₃), 0.85 (s, 9H, C(CH₃)₃), 0.17 (s, 3H, SiCH₃), 0.15 (s, 3H, SiCH₃), 0.06 (s, 9H, Si(CH₃)₃), 0.06 (s, 3H, SiCH₃), 0.03 (s, 3H, SiCH₃) *ppm*.

¹³C-NMR (101 MHz, Acetone-*d*₆): δ = 159.2, 158.6, 157.2, 153.7, 150.7, 147.6, 147.2, 139.5, 138.4, 133.7, 133.3, 131.5, 131.1, 129.9, 129.5, 129.5, 129.0, 128.5, 128.3, 127.2, 126.8, 126.3, 125.3, 125.2, 124.6, 122.2, 113.6, 112.2, 88.5, 85.3, 73.4, 71.0, 64.7, 63.8, 55.5, 26.4, 26.3, 19.0, 18.7, 18.0, -1.1, -4.3, -4.4, -5.0, -5.1 *ppm*.

HR-MS (ESI⁺): *m/z* calculated for [C₆₂H₈₁N₆O₆Si₃]⁺ ([M+H]⁺): 1089.5520, found: 1089.5520.

8-(2-aminopyrene)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-*N*²-(4,4'-dimethoxytrityl)-O⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine



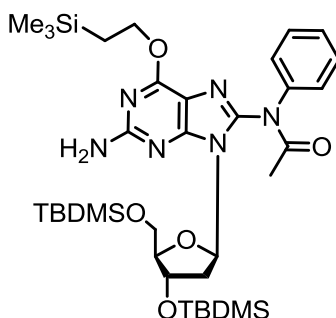
Yield: 58%

¹H-NMR (400 MHz, Acetone-*d*₆): δ = 8.67–8.59 (m, 2H, -C₁₆H₉), 8.25–7.96 (m, 8H, -C₁₆H₉, C8-NH), 7.45–7.38 (m, 2H, DMTr), 7.33–7.28 (m, 6H, DMTr), 7.26–7.18 (m, 1H, DMTr), 6.86–6.81 (m, 4H, DMTr), 6.39–6.30 (m, 1H, C1'-H), 6.29 (s, 1H, *N*²H), 4.78–4.69 (m, 1H, C3'-H), 4.16–4.11 (m, 1H, C4'-H), 4.04 (dd, 1H, C5'-H), 4.01–3.91 (m, 2H, O⁶-CH₂), 3.89 (dd, 1H, C5'-H), 3.77 (s, 6H, 2 x OCH₃), 3.40–3.18 (m, 1H, C2'-H), 2.33–2.21 (m, 1H, C2'-H), 0.92 (s, 9H, C(CH₃)₃), 0.90–0.83 (m, 2H, O⁶-CH₂-CH₂), 0.65 (s, 9H, C(CH₃)₃), 0.17 (s, 3H, SiCH₃), 0.17 (s, 3H, SiCH₃), -0.06 (s, 9H, Si(CH₃)₃), -0.13 (s, 3H, Si(CH₃)), -0.22 (s, 3H, Si(CH₃)) *ppm*.

^{13}C -NMR (100 MHz, Acetone- d_6): δ = 159.1, 157.0, 154.0, 149.1, 147.6, 139.5, 132.6, 132.1, 131.1, 129.9, 128.3, 128.3, 127.9, 127.3, 127.2, 127.0, 126.5, 126.4, 126.2, 125.7, 125.3, 125.2, 122.0, 113.5, 88.8, 85.9, 73.5, 70.9, 64.6, 63.8, 55.4, 39.3, 26.3, 26.2, 24.4, 18.8, 18.6, 18.0, 1.3, -4.3, -4.4, -5.3, -5.4 ppm.

HR-MS (ESI $^+$): m/z calculated for $[\text{C}_{64}\text{H}_{81}\text{N}_6\text{O}_6\text{Si}_3]^+$ ($[\text{M}+\text{H}]^+$): 1113.5520, found: 1113.5525.

8-(*N*-Acetyl-2-aminobenzene)-3',5'-bis-*O*-(*tert*-butyl-dimethylsilyl)-*O* 6 -(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine



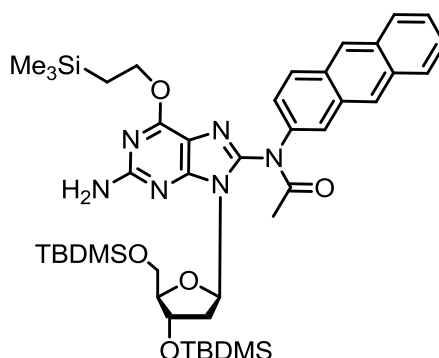
8-(2-aminobenzene)-3',5'-bis-*O*-(*tert*-butyl-dimethylsilyl)-*N* 2 -(4,4'-dimethoxytrityl)-*O* 6 -(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine (4.37 g, 4.42 mmol, 1.0 eq.) was dissolved in 50 mL dry pyridine. 4-Dimethylaminopyridine (0.27 g, 2.2 mmol, 0.5 eq.), trimethylamine (3.08 mL, 22.1 mmol, 5.0 eq.) and acetic anhydride (2.09 mL, 22.1 mmol, 5.0 eq.) were added and the reaction mixture was stirred at room temperature for 48 h under argon atmosphere. After evaporation to dryness the residue was dissolved in EtOAc (200 mL) and washed with water (2 x 60 mL). The organic phase was dried with MgSO_4 , filtered, concentrated under reduced pressure and dried for 3 h under high vacuum. Methanolic hydrochloric acid (30 mL, 0.01 M) was added and the mixture was stirred at room temperature for 20 h. After evaporation of the solvent, the residue was purified via column chromatography (isohexane/EtOAc 1:0 to 4:1; SiO_2). 8-(*N*-Acetyl-2-aminobenzene)-3',5'-bis-*O*-(*tert*-butyl-dimethylsilyl)-*O* 6 -(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine (2.23 g, 3.04 mmol, 69%) was obtained as a white yellowish foam.

^1H -NMR (400 MHz, Acetone- d_6): δ = 7.56–7.15 (m, 5H, $-\text{C}_6\text{H}_5$), 6.13 (t, 1H, C1'-H), 5.72 (s, br, 2H, NH_2), 4.80 (s, 1 H, C3'-H), 4.56 (t, 2H, $\text{O}^6\text{-CH}_2$), 3.94–3.84 (m, 2H, C5'-H, C4'-H), 3.77–3.72 (m, 1H, C5'-H), 3.44 (dt, 1H, C2'-H), 2.11–2.06 (m, 4H, C2'-H, COCH_3), 1.20 (t, 2H, $\text{O}^6\text{-CH}_2\text{-CH}_2$), 0.93 (s, 9H, $\text{C}(\text{CH}_3)_3$), 0.84 (s, 9H, $\text{C}(\text{CH}_3)_3$), 0.16 (s, 3H, SiCH_3), 0.13 (s, 3H, SiCH_3), 0.10 (s, 9H, $\text{Si}(\text{CH}_3)_3$), 0.01 (s, 3H, SiCH_3), -0.01 (s, 3H, SiCH_3) ppm.

^{13}C -NMR (201 MHz, Acetone- d_6): δ = 162.0, 159.3, 154.3, 144.2, 141.2, 130.0, 128.7, 128.3, 127.4, 114.2, 88.4, 84.8, 73.4, 64.8, 63.6, 55.4, 37.1, 26.2, 26.1, 18.8, 18.5, 17.9, -1.3, -4.5, -4.5, -5.2, -5.3 ppm.

HR-MS (ESI $^+$): calculated for $[\text{C}_{35}\text{H}_{61}\text{N}_6\text{O}_5\text{Si}_3]^+$ ($[\text{M}+\text{H}]^+$): 729.4006, found.: 729.4012.

8-(*N*-Acetyl-2-aminoanthracene)-3',5'-bis-*O*-(*tert*-butyl-dimethylsilyl)-*O*⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine



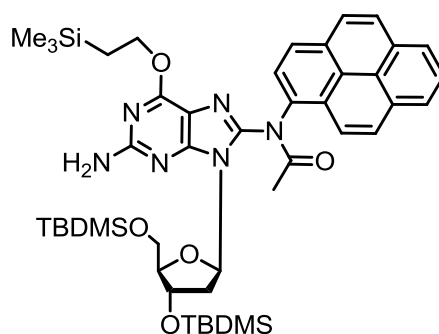
Yield: 74%

¹H-NMR (400 MHz, Acetone-*d*₆): δ = 8.56 (s, 1H, -C₁₄H₉), 8.49 (s, 1H, -C₁₄H₉), 8.12 (s, 1H, -C₁₄H₉), 8.06 (ddd, 3H, -C₁₄H₉), 7.64 (dd, 1H, -C₁₄H₉), 7.51 (dd, 2H, -C₁₄H₉), 6.21 (s, 1H, C1'-H), 5.74 (s, br, 2H, NH₂), 4.82–4.75 (m, 1H, C3'-H), 4.60 (t, 2H, O⁶-CH₂), 3.90–3.85 (m, 2H, C5'-H, C4'-H), 3.74 (d, 1H, C5'-H), 3.45 (ddd, 1H, C2'-H), 2.22–2.13 (m, 4H, C2'-H, COCH₃), 1.23 (t, 2H, O⁶-CH₂CH₂), 0.89 (s, 9H, C(CH₃)₃), 0.81 (s, 9H, C(CH₃)₃), 0.12 (s, 12H, 2x Si(CH₃)₂), -0.04 (s, 9H, Si(CH₃)₃) ppm.

¹³C-NMR (101 MHz, Acetone-*d*₆): δ = 162.1, 160.8, 159.3, 154.5, 144.3, 138.2, 133.0, 133.0, 132.3, 130.1, 129.8, 129.1, 128.9, 128.4, 127.3, 127.1, 126.7, 126.6, 114.4, 113.70, 88.5, 85.0, 73.5, 64.9, 63.9, 37.1, 26.2, 26.2, 23.3, 18.8, 18.5, 18.1, -1.3, -4.4, -4.5, -5.2, -5.3 ppm.

HR-MS (ESI⁺): *m/z* calculated for [C₄₃H₆₅N₆O₅Si₃]⁺ ([M+H]⁺): 829.4319, found: 829.4319.

8-(*N*-Acetyl-2-aminopyrene)-3',5'-bis-*O*-(*tert*-butyl-dimethylsilyl)-*O*⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine



Yield: 70%

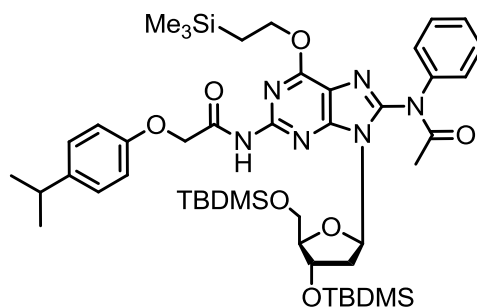
¹H-NMR (800 MHz, Acetone-*d*₆): δ = 8.71–8.62 (m, 1H, -C₁₆H₉), 8.47–7.92 (m, 8H, -C₁₆H₉), 6.64 (s, br, 2H, N²H), 6.54–6.40 (m, 1H, C1'-H), 4.81–4.72 (m, 1H, C3'-H), 4.65–4.53 (m, 1H, O⁶-CH₂), 4.49–4.29 (m, 1H, C5'-H), 4.12–3.81 (m, 3H, C4'-H, C5'-H), 3.67–3.54 (m, 1H, C2'-H), 2.40–2.34 (m, 1H, C2'-H), 1.99–1.87 (m, 3H, COCH₃), 1.27–1.13 (m, 2H, O⁶-CH₂-CH₂), 1.03–0.95 (m, 9H, C(CH₃)₃), 0.91–0.78 (m, 9H, C(CH₃)₃), 0.21 (s, 6H, Si(CH₃)₂), 0.15–0.05 (m, 9H, Si(CH₃)₃), 0.00–0.03 (m, 6H, Si(CH₃)₂) ppm.

¹³C-NMR (200 MHz, Acetone-*d*₆) δ = 159.1, 157.0, 154.0, 149.1, 147.6, 139.5, 132.6, 132.1, 131.1, 129.9, 128.3, 128.3, 127.9, 127.3, 127.2, 127.0, 126.5, 126.4, 126.2, 125.7, 125.3, 125.2, 122.0,

113.5, 88.8, 85.9, 73.5, 70.9, 64.6, 63.8, 55.4, 39.3, 26.3, 26.2, 24.4, 18.8, 18.6, 18.0, -1.3, -4.3, -4.4, -5.3, -5.4 ppm.

HR-MS (ESI⁺): m/z calculated for [C₄₅H₆₅N₆O₅Si₃]⁺ ([M+H]⁺): 853.4319, found: 853.4320.

8-(*N*-Acetyl-2-aminobenzene)-3',5'-bis-*O*-(*tert*-butyl-dimethylsilyl)-*N*²-isopropylphenoxyacetyl-*O*⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine



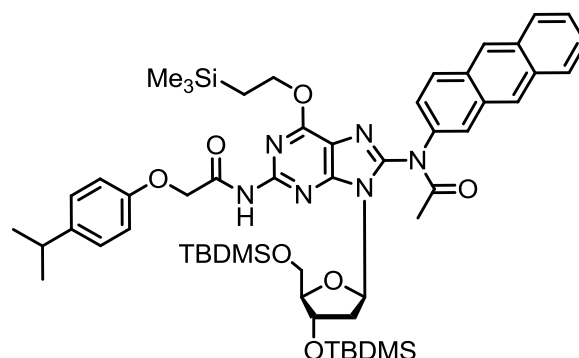
8-(*N*-Acetyl-2-aminobenzene)-3',5'-bis-*O*-(*tert*-butyl-dimethylsilyl)-*O*⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine (1.58 g, 2.17 mmol, 1.0 eq.) was dissolved in 50 mL dry pyridine, 4-isopropylphenoxyacetylchloride (1.50 mL, 8.67 mmol, 4.0 eq.) was added and the reaction mixture was stirred at room temperature for 24 h under argon atmosphere. After evaporation to dryness the residue was purified via column chromatography on silica gel (isohexane/EtOAc 1:0 to 7:1). 8-(*N*-Acetyl-2-aminobenzene)-3',5'-bis-*O*-(*tert*-butyl-dimethylsilyl)-*N*²-isopropylphenoxyacetyl-*O*⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine (1.70 g, 1.88 mmol, 86%) was obtained as a white yellowish foam.

¹H-NMR (400 MHz, Acetone-*d*₆): δ = 9.29 (s, *br*, 1H, *N*²H), 7.55–7.33 (m, 5H, -C₆H₅), 7.21 (d, 2H, *i*PrPac), 6.98 (d, 2H, *i*PrPac), 6.27 (t, 1H, C1'-H), 4.98 (s, 1H, C3'-H), 4.83 (s, 2H, *i*PrPac-CH₂), 4.68 (t, 2H, *O*⁶-CH₂), 3.99 (dd, 1H, C5'-H), 3.92–3.87 (m, 1H, C4'-H), 3.82 (dd, 1H, C5'-H), 3.44 (dt, 1H, C2'-H), 2.88 (hept, 1H, CH), 2.12 (s, 3H, COCH₃), 2.09–2.05 (m, 1H, C2'-H), 1.31–1.24 (m, 2H, *O*⁶-CH₂CH₂), 1.21 (d, 6H, CH(CH₃)₂), 0.92 (s, 9H, SiC(CH₃)₃), 0.81 (s, 9H, SiC(CH₃)₃), 0.16 (s, 3H, SiCH₃), 0.13 (s, 3H, SiCH₃), 0.11 (s, 9H, Si(CH₃)₃), -0.05 (s, 3H, SiCH₃), -0.05 (s, 3H, SiCH₃) ppm.

¹³C-NMR (101 MHz, Acetone-*d*₆): δ = 166.7, 161.8, 156.9, 153.1, 152.33, 147.2, 142.9, 130.2, 130.1, 128.8, 128.3, 128.2, 127.4, 117.8, 115.6, 113.6, 89.2, 73.4, 68.9, 66.2, 64.5, 55.5, 37.9, 34.1, 26.3, 26.2, 24.5, 18.8, 18.5, 18.0, -1.3, -4.5, -4.5, -5.1, -5.1 ppm.

HR-MS (ESI⁺): m/z calculated for [C₄₆H₇₃N₆O₇Si₃]⁺ ([M+H]⁺): 905.4843, found: 905.4851.

8-(*N*-Acetyl-2-aminoanthracene)-3',5'-bis-*O*-(*tert*-butyl-dimethylsilyl)-*N*²-isopropylphenoxyacetyl-*O*⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine



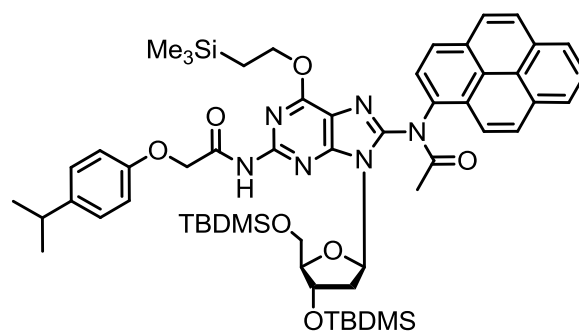
Yield: 70%

¹H-NMR (400 MHz, Acetone-*d*₆): δ = 9.30 (s, *br*, 1H, *N*²H), 8.58 (s, 1H, -C₁₄H₉), 8.51 (s, 1H, -C₁₄H₉), 8.15 (d, 1H, -C₁₄H₉), 8.07 (ddd, 3H, -C₁₄H₉), 7.67 (dd, 1H, -C₁₄H₉), 7.52 (dq, 2H, -C₁₄H₉), 7.20 (d, 2H, *i*PrPac), 6.97 (d, 2H, *i*PrPac), 6.36 (t, 1H, C1'-H), 4.97 (s, 1H, C3'-H), 4.83 (s, 2H, *i*PrPac-CH₂), 4.70 (t, 2H, O⁶-CH₂), 3.97 (dd, 1H, C5'-H), 3.92-3.88 (m, 1H, C4'-H), 3.81 (d, 1H, C5'-H), 3.46 (dt, 1H, C2'-H), 2.87 (hept, 1H, CH), 2.22 (s, 3H, OCH₃), 2.09-2.06 (m, 1H, C2'-H), 1.27 (t, 2H, O⁶-CH₂CH₂), 1.20 (d, 6H, CH(CH₃)₂), 0.89 (s, 9H, C(CH₃)₃), 0.78 (s, 9H, C(CH₃)₃), 0.13 (s, 3H, SiCH₃), 0.12 (s, 9H, Si(CH₃)₃), 0.07 (s, 3H, SiCH₃), -0.08 (s, 3H, SiCH₃), -0.08 (s, 3H, SiCH₃) *ppm*.

¹³C-NMR (101 MHz, Acetone-*d*₆): δ = 166.7, 161.8, 159.3, 156.9, 153.2, 152.4, 147.2, 142.9, 138.1, 133.1, 132.2, 131.2, 130.4, 129.8, 129.1, 128.9, 128.4, 128.2, 127.5, 127.2, 126.8, 126.7, 117.9, 115.6, 113.7, 89.2, 85.3, 73.3, 68.9, 66.3, 64.4, 37.9, 34.1, 26.2, 26.2, 24.5, 23.3, 18.8, 18.5, 18.1, -1.3, -4.5, -4.5, -5.1, -5.2 *ppm*.

HR-MS (ESI⁺): *m/z* calculated for [C₅₄H₇₇N₆O₇Si₃]⁺ ([M+H]⁺): 1005.5156, found: 1005.5158.

8-(*N*-Acetyl-2-aminopyrene)-3',5'-bis-*O*-(*tert*-butyl-dimethylsilyl)-*N*²-isopropylphenoxyacetyl-*O*⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine



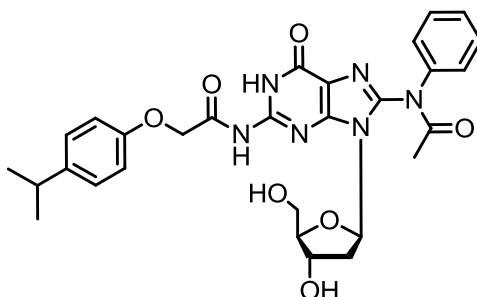
Yield: 92%

¹H-NMR (400 MHz, Acetone-*d*₆): δ = 9.26 (s, 1H, *N*²H), 8.46-8.02 (m, 9H, -C₁₆H₉), 7.24-7.15 (d, 2H, *i*PrPac), 6.97 (d, 2H, *i*PrPac), 6.68-6.52 (m, 1H, C1'-H), 5.25-4.92 (m, 1H, C3'-H), 4.82 (s, 2H, *i*PrPac-CH₂), 4.72-4.39 (m, 2H, O⁶-CH₂), 4.13-3.97 (m, 2H, C4'-H, C5'-H), 3.97-3.82 (m, 1H, C5'-H), 3.65-3.46 (m, 1H, C2'-H), 2.86 (hept, 1H, CH), 2.81 (s, 3H, COCH₃), 2.79-2.76 (m, 1H, C2'-H), 1.21 (s, 3H, CH(CH₃)₂), 1.19 (s, 3H, CH(CH₃)₂), 1.15-1.11 (m, 2H, O⁶-CH₂-CH₂), 0.94 (s, 9H, C(CH₃)₃), 0.82 (s, 9H, C(CH₃)₃), 0.16 (m, 6H, Si(CH₃)₂), 0.04 (s, 9H, Si(CH₃)₃), -0.04 (s, 6H, Si(CH₃)₂) *ppm*.

^{13}C -NMR (100 MHz, Acetone- d_6): δ = 166.7, 156.9, 153.2, 142.9, 132.1, 128.2, 128.0, 127.6, 127.0, 126.2, 115.7, 115.6, 89.4, 85.6, 73.5, 68.9, 66.1, 34.0, 26.3, 26.2, 24.5, 21.4, 18.9, 18.5, 18.0, -1.3, -4.4, -5.1 ppm.

HR-MS (ESI $^+$): m/z calculated for $[\text{C}_{56}\text{H}_{77}\text{N}_6\text{O}_7\text{Si}_3]^+$ $[\text{M}+\text{H}]^+$: 1029.5156, found: 1029.5156.

8-(*N*-Acetyl-2-aminobenzene)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine



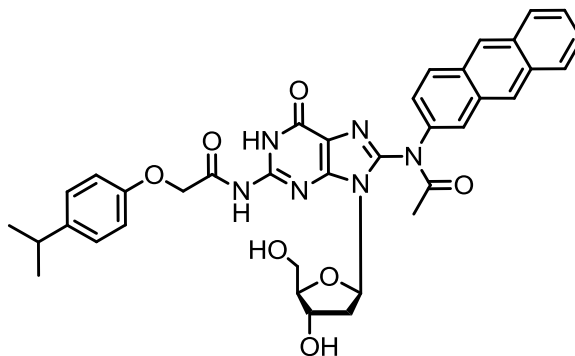
8-(*N*-Acetyl-2-aminobenzene)-3',5'-bis-*O*-(*tert*-butyl-dimethylsilyl)-*N*²-isopropylphenoxyacetyl-*O*⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine (1.10 g, 1.21 mmol, 1.0 eq.) was dissolved in dry THF (60 mL), 1 M TBAF in THF (4.86 mL, 4.0 eq.) and acetic acid (0.42 mL, 7.3 mmol, 6.0 eq.) were added and the reaction mixture was stirred at room temperature for 24 h under argon atmosphere. After evaporation to dryness the residue was purified via column chromatography on silica gel (DCM/MeOH 1:0 to 40:1). 8-(*N*-Acetyl-2-aminobenzene)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine (0.55 g, 0.95 mmol, 78%) was obtained as a yellowish foam.

^1H -NMR (400 MHz, DMSO- d_6): δ = 11.79 (s, *br*, 1H, *N*¹-H), 11.38 (s, *br*, 1H, *N*²-H), 7.64 (dd, 2H, C_6H_5), 7.29 (dd, 2H, $-\text{C}_6\text{H}_5$), 7.18 (d, 2H, *i*PrPac), 6.94 (t, 1H, $-\text{C}_6\text{H}_5$), 6.91 (d, 2H, *i*PrPac), 6.38 (t, 1H, $\text{C}1'$ -H), 5.39 (d, *br*, 1H, $\text{C}3'$ -OH), 4.84 (s, 2H, *i*PrPac- CH_2), 4.60 (dd, 1H, $\text{C}3'$ -H), 4.38 (dd, 1H, $\text{C}5'$ -H), 4.20 (dd, 1H, $\text{C}5'$ -H), 3.97 (dt, 1H, $\text{C}4'$ -H), 3.12 (dt, 1H, $\text{C}2'$ -H), 2.84 (hept, 1H, CH), 2.23 (ddd, 1H, $\text{C}2'$ -H), 1.97 (s, 3 H, COCH_3), 1.17 (d, 6H, $\text{CH}(\text{CH}_3)_2$) ppm.

^{13}C -NMR (101 MHz, DMSO- d_6): δ = 170.2, 155.7, 153.8, 145.0, 141.3, 140.7, 127.2, 121.1, 117.6, 114.4, 83.9, 82.8, 70.6, 69.6, 66.5, 64.3, 37.1, 32.6, 24.1 ppm.

HR-MS (ESI $^+$): m/z calculated for $[\text{C}_{29}\text{H}_{33}\text{N}_6\text{O}_7]^+$ $[\text{M}+\text{H}]^+$: 577.2405, found: 577.2413.

8-(*N*-Acetyl-2-aminoanthracene)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine



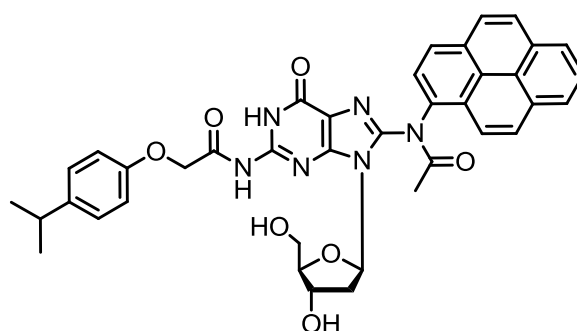
Yield: 76%

^1H -NMR (400 MHz, Acetone- d_6): δ = 11.83 (s, *br*, 1H, $N^1\text{H}$), 10.07 (s, *br*, 1H, $N^2\text{H}$), 8.58-8.52 (m, 2H, -C₁₄H₉), 8.15 (d, 1H, -C₁₄H₉), 8.10-8.04 (m, 3H, -C₁₄H₉), 7.67 (dd, 1H, -C₁₄H₉), 7.52 (dd, 2H, -C₁₄H₉), 7.21 (d, 2H, *i*PrPac), 7.01 (d, 2H, *i*PrPac), 6.39 (t, 1H, C1'-H), 5.97 (s, *br*, 1H, C3'-OH), 4.86 (s, 2H, *i*PrPac-CH₂), 4.61 (s, *br*, 1H, C5'-OH), 4.43-4.35 (m, 1-H, C3'-H), 4.08-4.01 (m, 1H, C4'-H), 3.79 (d, 1H, C5'-H), 3.74 (s, 1H, C5'-H), 3.28-3.18 (m, 1H, C2'-H), 2.91-2.84 (m, 1H, CH), 2.20 (s, 3H, OCH₃), 2.10-2.05 (m, 1H, C2'-H), 1.20 (d, 6H, CH(CH₃)₂) *ppm*.

^{13}C -NMR (101 MHz, Acetone- d_6): δ = 172.0, 159.5, 156.4, 155.1, 143.4, 138.3, 133.1, 132.2, 130.5, 129.5, 129.1, 128.9, 128.8, 128.3, 128.1, 127.5, 127.2, 126.8, 126.8, 121.1, 115.8, 89.7, 86.5, 72.8, 71.1, 67.9, 65.8, 63.6, 39.1, 34.1, 24.5, 23.2 *ppm*.

HR-MS (ESI⁺): *m/z* calculated for [C₃₇H₃₇N₆O₇]⁺ ([M+H]⁺): 677.2718, found: 677.2726.

8-(*N*-Acetyl-2-aminopyrene)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine



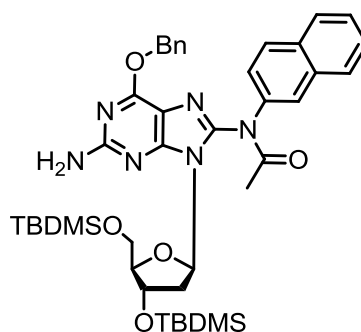
Yield 90%

^1H -NMR (400 MHz, DMSO- d_6): δ = 9.75 (s, *br*, 1H, $N^1\text{H}$), 9.08 (*br s*, 1H, $N^2\text{H}$), 8.95-8.50 (m, 9H, -C₁₆H₉), 7.63 (d, 2H, *i*PrPac), 7.43 (d, 2H, *i*PrPac), 7.20-7.00 (m, 1H, C1'-H), 5.60-5.44 (m, 1H, C3'-H), 5.33 (s, 2H, *i*PrPac-CH₂), 5.09-4.98 (m, 1H, C3'-OH), 4.69-3.96 (m, 3H, C4'-H, C5'-H), 3.75-3.55 (m, 1H, C2'-H), 2.99-2.83 (m, 1H, CH), 2.41 (s, 3H, COCH₃), 2.30-2.16 (m, 1H, C2'-H), 1.64 (s, 3H, CH(CH₃)₂), 1.63 *ppm* (s, 3H, CH(CH₃)₂) *ppm*.

^{13}C NMR (100 MHz, DMSO- d_6): δ = 166.7, 152.4, 141.7, 137.8, 137.8, 137.7, 136.0, 134.7, 125.3, 125.3, 78.4, 77.5, 76.1, 69.1, 43.7, 40.1, 39.9, 39.7, 39.5, 39.3, 39.2, 38.9, 34.2, 34.1, 34.1, 30.2, 30.1, 27.7, 23.5, 8.3 *ppm*.

HR-MS (ESI⁻): *m/z* calculated for [C₃₉H₃₅N₆O₇]⁻ ([M-H]⁻): 699.2573, found: 699.2579.

8-(*N*-Acetyl-2-aminonaphthalene)-3',5'-bis-*O*-(*tert*-butyl-dimethylsilyl)-*O*⁶-benzyl-2'-deoxyguanosine



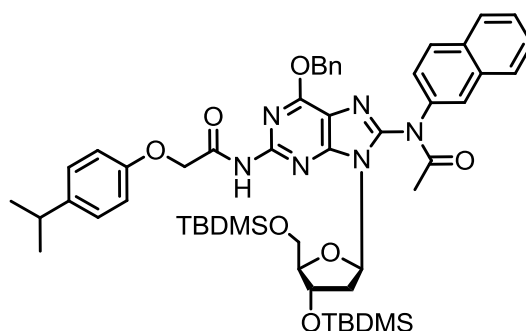
8-(2-aminonaphthalene)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-*N*²-(4,4'-dimethoxytrityl)-*O*⁶-benzyl-ethyl-2'-deoxyguanosine^[14] (6.47 g, 6.28 mmol, 1.0 eq.) was dissolved in 120 mL dry pyridine. 4-Dimethylaminopyridine (0.38 g, 3.14 mmol, 0.5 eq.), trimethylamine (4.38 mL, 31.4 mmol, 5.0 eq.) and acetic anhydride (2.97 mL, 31.4 mmol, 5.0 eq.) were added and the reaction mixture was stirred at room temperature for 48 h under argon atmosphere. After evaporation to dryness the residue was dissolved in EtOAc (200 mL) and washed with water (2 x 60 mL). The organic phase was dried with MgSO₄, filtered, concentrated under reduced pressure and dried for 3 h under high vacuum. Methanolic hydrochloric acid (30 mL, 0.01 M) was added and the mixture was stirred at r.t. for 20 h. After evaporation of the solvent, the residue was purified via column chromatography (*iso*hexane/EtOAc 1:0 to 7:1; SiO₂). 8-(*N*-Acetyl-2-aminonaphthalene)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-*O*⁶-benzyl-2'-deoxyguanosine (3.99 g, 5.19 mmol, 83%) was obtained as a white brown-yellowish foam.

¹H-NMR (400 MHz, Acetone-*d*₆): δ = 7.95-7.84 (m, 4H, -C₁₀H₇), 7.65-7.47 (m, 5H, 3x -C₁₀H₇, 2x Bn), 7.42-7.28 (m, 3H, Bn), 6.21 (s, 1H, C1'-H), 5.84 (s, *br*, 1H, *N*²H), 5.53 (s, *br*, 2H, Bn-CH₂), 4.81 (s, 1H, C3'-H), 3.94-3.82 (m, 2H, C4'-H, C5'-H), 3.80-3.70 (m, 1H, C5'-H), 3.46 (dd, 2H, C2'-H), 2.14 (s, 3H, COCH₃), 0.91 (s, 9H, C(CH₃)₃), 0.82 (s, 9H, C(CH₃)₃), 0.17-0.08 (m, 6H, Si(CH₃)₂), -0.01-(-0.04) (m, 6H, Si(CH₃)₂) *ppm*.

¹³C-NMR (101 MHz, Acetone-*d*₆): δ = 161.71, 160.64, 154.74, 137.72, 134.40, 129.44, 129.24, 128.89, 128.83, 128.49, 127.42, 114.23, 88.50, 84.96, 73.43, 68.29, 63.84, 41.83, 37.08, 26.23, 26.18, 20.83, 18.79, 18.51, -4.45, -5.21 *ppm*.

HR-MS (ESI⁺): *m/z* calculated for [C₄₁H₅₇N₆O₅ Si₂]⁺ ([M+H]⁺): 769.3851, found.: 769.39207.

8-(*N*-Acetyl-2-naphthalene)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-*N*²-isopropylphenoxyacetyl-*O*⁶-benzyl-2'-deoxyguanosine



8-(*N*-Acetyl-2-aminonaphthalene)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-*O*⁶-benzyl-2'-deoxyguanosine (3.99 g, 5.19 mmol, 1.0 eq.) was dissolved in 120 mL dry pyridine, 4-*isopropylphenoxyacetylchloride* (3.6 mL, 20.7 mmol, 4.0 eq.) was added and the reaction mixture was stirred at room temperature for 24 h under argon atmosphere. After evaporation to dryness the residue was purified via column chromatography on silica gel (*iso*hexane/EtOAc 1:0 to 6:1). 8-(*N*-Acetyl-2-naphthalene)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-*N*²-*isopropylphenoxyacetyl-O*⁶-benzyl-2'-deoxyguanosine (4.54 g, 4.80 mmol, 92%) was obtained as a white yellowish foam.

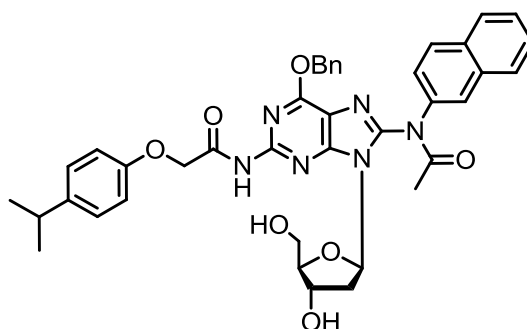
¹H-NMR (400 MHz, Acetone-*d*₆): δ = 9.44 (s, 1H, *N*²H), 8.00-7.85 (m, 4H, -C₁₀H₇), 7.68-7.50 (m, 5H, 3x -C₁₀H₇, 2x Bn), 7.41-7.29 (m, 3H, Bn), 7.23-7.18 (m, 2H, *i*PrPac), 7.01-6.95 (m, 2H, *i*PrPac), 6.40-6.31 (m, *br*, 1H, C1'-H), 5.68-5.55 (m, 2H, Bn-CH₂), 5.03-4.94 (m, 1H, C3'-H), 4.84 (s, 2H, *i*PrPac-CH₂), 4.00-3.78 (m, *br*, 3H, 2x C5'-H, C4'-H), 3.43-3.31 (m, 2H, C2'-H₂), 2.93-2.76 (m, 1H, CH), 2.17

(s, 3H, COCH₃), 1.21 (d, 3H, CH(CH₃)₂), 1.16 (d, 3H, CH(CH₃)₂), 0.91 (s, 9H, C(CH₃)₃), 0.78 (s, 9H, C(CH₃)₃), 0.19-0.05 (m, 6H, Si(CH₃)₂), -0.07 (s, 6H, Si(CH₃)₂) ppm.

¹³C-NMR (101 MHz, Acetone-*d*₆) δ = 170.46, 166.74, 161.40, 156.95, 153.34, 152.22, 147.50, 142.91, 137.21, 134.44, 129.61, 129.33, 129.11, 128.90, 128.56, 128.20, 117.73, 115.64, 89.21, 85.31, 73.27, 68.91, 64.41, 58.97, 37.96, 34.05, 26.23, 26.20, 24.50, 21.10, 18.80, 18.45, -4.47, -5.15 ppm.

HR-MS (ESI⁻): m/z calculated for [C₅₂H₆₇N₆O₇Si₂]⁻ ([M-H]⁻): 943.4688, found: 943.46036.

8-(*N*-Acetyl-2-aminonaphthalene)-*N*²-isopropylphenoxyacetyl-*O*⁶-benzyl-2'-deoxyguanosine



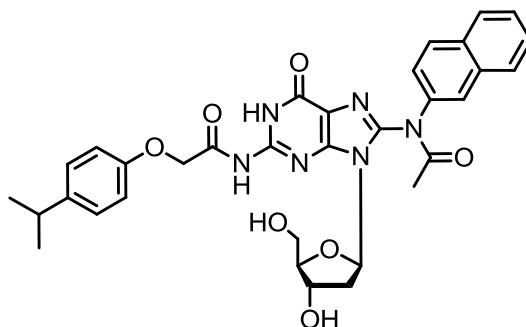
8-(*N*-Acetyl-2-naphthalene)-3',5'-bis-*O*-(*tert*-butyl-dimethylsilyl)-*N*²-isopropylphenoxyacetyl-*O*⁶-benzyl-2'-deoxyguanosine (4.50 g, 4.76 mmol, 1.0 eq.) was dissolved in dry THF (120 mL), 1 M TBAF in THF (19.0 mL, 4.0 eq.) and acetic acid (1.63 mL, 28.6 mmol, 6.0 eq.) were added and the reaction mixture was stirred at room temperature for 24 h under argon atmosphere. After evaporation to dryness the residue was purified via column chromatography on silica gel (DCM/MeOH 1:0 to 40:1). 8-(*N*-Acetyl-2-aminonaphthalene)-*N*²-isopropylphenoxyacetyl-*O*⁶-benzyl-2'-deoxyguanosine (3.10 g, 4.32 mmol, 91%) was obtained as a yellowish foam.

¹H-NMR (400 MHz, Acetone-*d*₆): δ = 9.44 (s, 1H, *N*²-H), 8.03-7.87 (m, 4H, -C₁₀H₇), 7.71-7.50 (m, 5H, 3x -C₁₀H₇, 2x Bn), 7.41-7.30 (m, 3H, Bn), 7.22-7.14 (m, 2H, *i*PrPac), 7.02-6.96 (m, 2H, *i*PrPac), 6.43 (t, 1H, C1'-H), 5.69-5.57 (s, 2H, Bn-CH₂), 4.88 (s, 2H, *i*PrPac-CH₂), 4.80 (s, *br*, 1H, C3'-H) 4.04 (s, *br*, 1H, C4'-H), 3.90-3.71 (m, 2H, C5'-H), 3.30-2.90 (m, 2H, C2'-H) 2.86 (hept, 1H, CH), 1.96 (s, 3H, COCH₃), 1.20 (d, 6H, CH(CH₃)₂) ppm.

¹³C-NMR (101 MHz, Acetone-*d*₆): δ = 172.01, 167.33, 161.47, 156.89, 153.05, 152.11, 147.28, 142.80, 137.10, 134.41, 129.77, 129.30, 129.13, 128.92, 128.55, 128.15, 127.63, 117.75, 115.63, 89.85, 86.39, 72.60, 69.50, 68.71, 63.54, 39.13, 34.02, 24.49, 20.50 ppm.

HR-MS (ESI⁺): m/z calculated for [C₄₀H₄₁N₆O₆]⁺ ([M+H]⁺): 717.2958, found: 717.30365.

8-(*N*-Acetyl-2-aminonaphthalene)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine



8-(*N*-Acetyl-2-aminonaphthalene)-*N*²-isopropylphenoxyacetyl-*O*⁶-benzyl-2'-deoxyguanosine (0.50 g, 0.69 mmol, 1.0 eq.) was dissolved in 25 mL MeOH/THF 1:1, Pd/C (0.05g) was added and argon was exchanged with hydrogen atmosphere. The reaction mixture was stirred at room temperature for 20 h. After evaporation to dryness the residue was purified via column chromatography on silica gel (DCM/MeOH 1:0 to 20:1). 8-(*N*-Acetyl-2-aminonaphthalene)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine (0.40 g, 0.63 mmol, 90%) was obtained as a white solid.

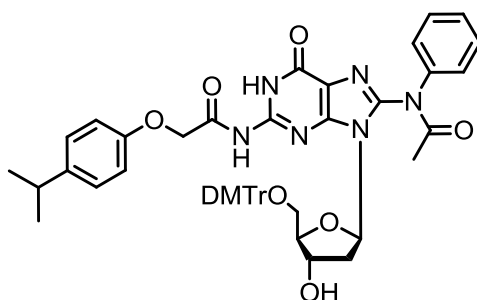
¹H-NMR (400 MHz, DMSO-*d*₆): δ = 11.92 (s, *br*, 2H, *N*¹-H, *N*²-H), 8.00 (d, 1H, -C₆H₅), 7.97-7.91 (m, 3H, -C₆H₅), 7.63-7.50 (m, 2H, -C₆H₅), 7.45-7.33 (m, 1H, -C₆H₅), 7.16 (d, 2H, *i*PrPac), 6.88 (d, 2H, *i*PrPac), 6.31-6.13 (m, 1H, C1'-H), 5.24 (s, *br*, 1H, C3'-OH), 4.93 (s, *br*, 1H, C5'-OH), 4.81 (s, 2H, *i*PrPac-CH₂), 4.51-4.40 (m, 1H, C3'-H), 3.90-3.80 (m, 1H, C4'-H), 3.68-3.61 (m, 1H, C5'-H), 3.56-3.47 (m, 1H, C5'-H), 3.36-3.29 (m, 2H, C2'-H), 2.80 (sept, 1H, CH), 2.10 (s, 3H, COCH₃), 1.17 (d, 3H, CH(CH₃)₂), 1.15 (d, 3H, CH(CH₃)₂) *ppm*.

¹³C-NMR (100 MHz, DMSO-*d*₆): δ = 169.3, 155.9, 149.6, 141.1, 136.1, 132.9, 129.0, 127.9, 127.6, 127.1, 127.1, 126.8, 124.9, 123.9, 118.8, 114.4, 114.2, 88.1, 84.0, 70.9, 67.0, 64.6, 62.0, 51.7, 36.9, 34.4, 32.6, 30.4, 24.1, 22.7 *ppm*.

HR-MS (ESI⁻): *m/z* calculated for [C₃₃H₃₂N₆O₇]⁻ ([M-H]⁻): 626.6591, found: 625.2417.

The DMT protection and the preparation of the phosphoramidites of the phenyl, naphthyl, pyrenyl and the anthryl adducts were carried out analogous.

8-(*N*-Acetyl-2-aminobenzene)-5'-*O*-(4,4'-dimethoxytrityl)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine



8-(*N*-Acetyl-2-aminobenzene)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine (200 mg, 0.347 mmol, 1.0 eq.) was evaporated two times from dry pyridine and then dissolved in 2 mL dry pyridine. Molecular sieves (4 Å) were added and the solution was stirred at room temperature for 2 h under argon atmosphere. 4,4'-dimethoxytritylchloride (153 mg,

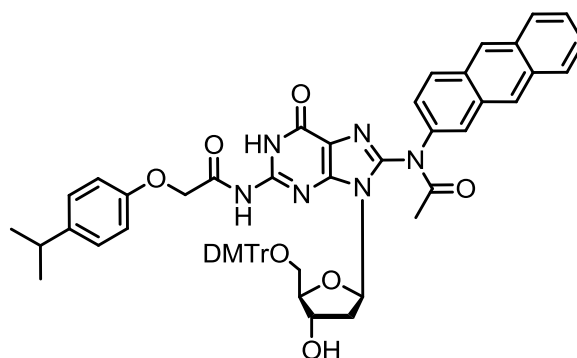
0.451 mmol, 1.3 eq.) was added and the reaction mixture was stirred at room temperature for 3 h. After evaporation to dryness the residue was purified via column chromatography on silica gel (DCM/MeOH 1:0 to 10:1 + 1% pyridine). 8-(*N*-Acetyl-2-aminobenzene)-5'-O-(4,4'-dimethoxytrityl)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine (274 mg, 0.312 mmol, 90%) was obtained as a yellow foam.

¹H-NMR (599 MHz, DMSO-*d*₆): δ = 11.74 (s, 1H, *N*¹-H), 11.31 (s, 1H, *N*²-H), 7.44-7.36 (m, 4H, -C₆H₅), 7.33-7.29 (m, 1H, -C₆H₅), 7.27-7.22 (m, 2H, DMTr), 7.16-7.14 (m, 2H, *i*PrPac), 7.14-7.02 (m, 7H, DMTr), 6.86 (d, 2H, *i*PrPac), 6.72-6.65 (m, 2H, DMTr), 6.66-6.54 (m, 2H, DMTr), 6.24 (s, 1H, C1'-H), 5.17 (d, 1H, C3'-OH), 4.82 (d, 1H, CH₂), 4.70 (d, 1H, *i*PrPacCH₂), 4.59 (m, 1H, C3'-H), 4.00 (s, 1H, C4'-H), 3.65 (s, 3H, OCH₃), 3.62 (s, 3H, OCH₃), 3.37 (s, *br*, 1H, C5'-H), 3.10 (d, 1H, C5'-H), 2.81 (hept, 1H, CH), 2.04 (m, 4H, C2'-H, COCH₃), 1.14 (d, 6H, CH(CH₃)₂) *ppm*.

¹³C-NMR (151 MHz, DMSO-*d*₆): δ = 171.0, 170.5, 158.6, 158.3, 158.2, 156.1, 154.8, 147.8, 147.0, 145.3, 143.4, 141.9, 136.2, 136.0, 130.2, 130.1, 129.8, 128.2, 127.8, 127.7, 126.8, 119.6, 114.9, 113.2, 113.1, 87.2, 85.1, 84.2, 71.3, 67.0, 65.4, 55.3, 33.0, 24.5, 23.1 *ppm*.

HR-MS (ESI⁺): *m/z* calculated for [C₅₀H₅₁N₆O₉]⁺ ([M+H]⁺): 879.3712, found: 879.3723.

8-(*N*-Acetyl-2-aminoanthracene)-5'-O-(4,4'-dimethoxytrityl)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine



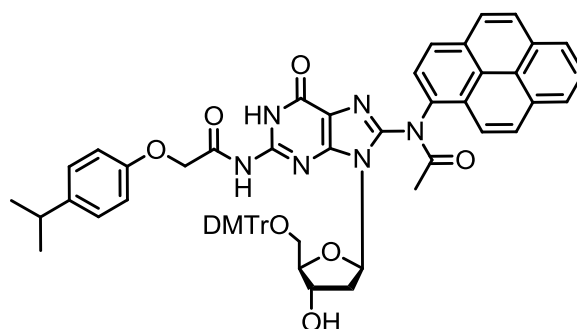
Yield: 64%

¹H-NMR (599 MHz, DMSO-*d*₆): δ = 11.72 (s, 1H, *N*¹H), 11.30 (s, 1H, *N*²H), 8.62-8.46 (m, 2H, -C₁₄H₉), 8.07 (dd, 2H, -C₁₄H₉), 7.78 (dd, 2H, -C₁₄H₉), 7.56 (s, 1H, -C₁₄H₉), 7.53-7.42 (m, 1H, -C₁₄H₉), 7.35 (m, 1H, -C₁₄H₉), 7.33-7.23 (m, 3H, DMTr), 7.18-7.17 (m, 2H, DMTr), 7.16-7.13 (m, 2H, *i*PrPac), 7.13-7.01 (m, 6H, DMTr), 6.85 (d, 2H, *i*PrPac), 6.74-6.47 (m, 2H, DMTr), 6.24 (s, 1H, C1'-H), 5.17 (s, 1H, C3'-OH), 4.81 (dd, 1H, CH₂), 4.69 (d, 1H, *i*PrPacCH₂), 4.55 (s, 1H, C3'-H), 4.10-4.00 (m, 2H, C5'-H, C4'-H), 3.87 (s, 1H, C5'-H), 3.64 (s, 6H, 2 x OCH₃), 3.18-3.08 (td, 1H, C2'-H), 2.80 (hept, 1H, CH), 2.20-2.00 (m, 4H, C2'-H, COCH₃), 1.14 (d, 6H, CH(CH₃)₂) *ppm*.

¹³C-NMR (151 MHz, DMSO-*d*₆): δ = 171.0, 158.2, 156.1, 154.8, 145, 141.9, 136.0, 131.9, 131.1, 130.1, 128.5, 128.2, 127.8, 127.7, 126.8, 126.4, 119.6, 114.9, 113.2, 113.0, 87.3, 85.6, 84.2, 71.3, 67.0, 65.5, 55.3, 55.2, 33.0, 24.5, 23.1 *ppm*.

HR-MS (ESI⁻): *m/z* calculated for [C₅₈H₅₃N₆O₉]⁻ ([M-H]⁻): 977.3879, found: 977.3882.

8-(*N*-Acetyl-2-aminopyrene)-5'-*O*-(4,4'-dimethoxytrityl)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine

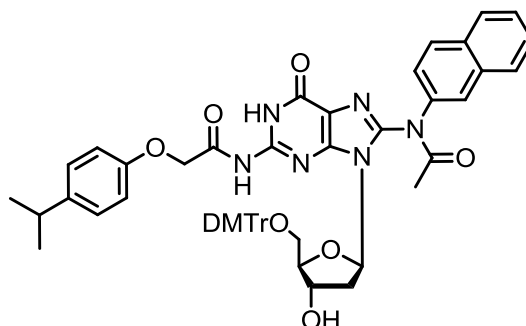


Yield: 84%

¹H-NMR (800 MHz, Acetone-*d*₆): δ = 11.65 (s, 1H, *N*¹H), 10.24 (s, 1H, *N*²H), 8.82-8.03 (m, 9H, -C₁₆H₉), 7.37-7.32 (m, 2H, DMTr), 7.33-7.23 (m, 5H, DMTr), 7.22-7.18 (m, 2H, *i*PrPac), 7.15-7.04 (m, 4H, DMTr), 6.99-6.90 (m, 2H, *i*PrPac), 6.76-6.65 (m, 2H, DMTr), 6.46-6.32 (m, 1H, C1'-H), 5.19 (s, 1H, C3'-OH), 4.86 (s, 1H, C3'-H), 4.82-4.72 (m, 2H, *i*PrPacCH₂), 4.68-4.26 (m, 3H, C4'-H, C5'-H), 3.69-3.60 (m, 6H, OCH₃), 3.29-3.22 (m, 1H, C2'-H), 2.83-2.76 (hept, 1H, CH), 2.00-1.95 (m, 1H, C2'-H), 1.29 (s, 3H, COCH₃), 1.20 ppm (s, 6H, 2x CH(CH₃)₂) *ppm*.

HR-MS (ESI⁺): *m/z* calculated for [C₆₀H₅₃N₆O₉]⁺ ([M-H]⁺): 1001.3879, found: 1001.3887.

8-(*N*-Acetyl-2-aminonaphthalene)-5'-*O*-(4,4'-dimethoxytrityl)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine



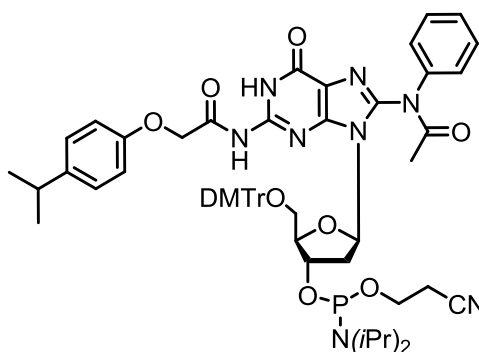
Yield: 65%

¹H NMR (599 MHz, DMSO-*d*₆) δ 11.73 (s, 1H, *N*¹H), 11.30 (s, 1H, *N*²H), 7.93 (d, 1H, -C₁₀H₇), 7.90 (d, 1H, -C₁₀H₇), 7.81 (s, 3H, -C₁₀H₇), 7.50 (t, 1H, -C₁₀H₇), 7.45 (s, 1H, -C₁₀H₇), 7.24 (d, 2H, DMTr), 7.15 (d, 2H, *i*PrPac), 7.13-6.92 (m, 7H, DMTr), 6.85 (d, 2H, *i*PrPac), 6.62 (d, 4H, DMTr), 6.31 (s, 1H, C1'-H), 5.15 (s, 1H, C3'-OH), 4.81 (d, 1H, CH₂), 4.69 (d, 1H, *i*PrPacCH₂), 4.54 (s, 1H, C3'-H), 4.02 (s, 1H, C4'-H), 3.63 (s, 3H, OCH₃), 3.61 (s, 3H, OCH₃), 3.37 (s, 1H, C5'-H), 3.12 (m, 1H, C5'-H), 3.10-3.00 (m, 1H, C2'-H), 2.93-2.75 (hept, 1H, CH) 2.20-2.06 (s, 4H, COCH₃, C2'-H), 1.14 (d, 6H, CH(CH₃)₂) *ppm*.

¹³C NMR (151 MHz, DMSO-*d*₆) δ 171.0, 158.2, 156.1, 154.8, 145.3, 141.9, 136.0, 133.4, 132.0, 130.2, 130.1, 128.2, 128.0, 127.8, 127.7, 127.2, 126.8, 119.8, 114.8, 113.2, 113.0, 87.3, 85.6, 84.3, 71.3, 67.0, 65.7, 55.3, 33.0, 24.5, 23.9 *ppm*.

HR-MS (ESI⁺): *m/z* calculated for [C₅₄H₅₁N₆O₉]⁺ ([M-H]⁺): 927.3723, found: 927.3724.

8-(*N*-Acetyl-2-aminobenzene)-3'-O-(2'-cyanoethoxydiisopropylamino-phosphino)-5'-O-(4,4'-dimethoxytrityl)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine (1)

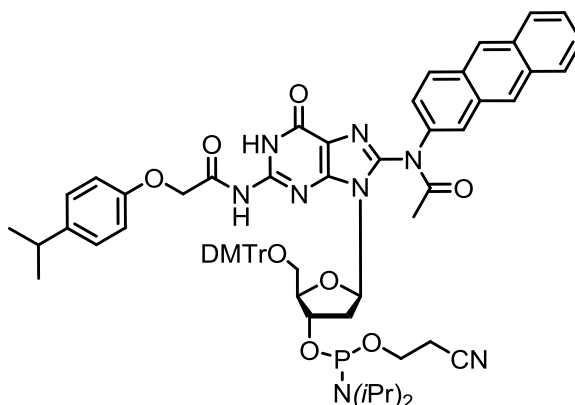


8-(*N*-Acetyl-2-aminobenzene)-5'-O-(4,4'-dimethoxytrityl)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine (200 mg, 0.227 mmol, 1.0 eq.) and Diisopropylammoniumtetrazolide (144 mg, 0.83 mmol, 3.6 eq.) were dissolved in dry DCM (2 mL) and degassed three times. 3-((bis(diisopropylamino)phosphino)oxy)propanenitrile (146 μ L, 0.46 mmol, 2.0 eq.) was added and the mixture was stirred at room temperature for 2 h. 8-(*N*-Acetyl-2-aminobenzene)-3'-O-(2'-cyanoethoxydiisopropylamino-phosphino)-5'-O-(4,4'-dimethoxytrityl)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine **1** (245 mg, 0.226 mmol, 99%) was obtained as a yellow solid.

³¹P-NMR (162 MHz, CDCl₃): δ = 148.9 ppm.

HR-MS (ESI⁺): m/z calculated for [C₅₉H₆₆N₈O₁₀P]⁺ ([M-H]⁺): 1077.4645, found: 1077.4647.

8-(*N*-Acetyl-2-aminoanthracene)-3'-O-(2'-cyanoethoxydiisopropylamino-phosphino)-5'-O-(4,4'-dimethoxytrityl)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine (3)

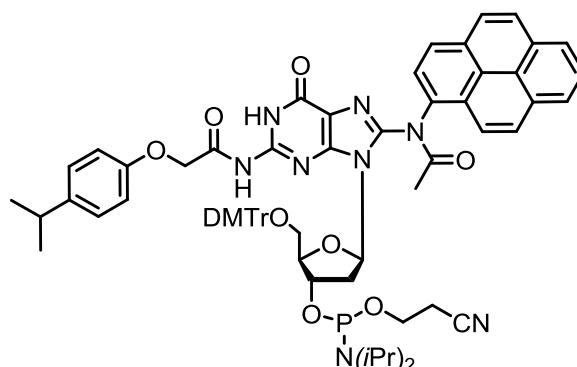


Yield: 99%

³¹P-NMR (162 MHz, CDCl₃): δ = 149.3 ppm.

HR-MS (ESI⁺): m/z calculated for [C₆₇H₇₀N₈O₁₀P]⁺ ([M-H]⁺): 1177.4958, found: 1177.4960.

8-(*N*-Acetyl-2-aminopyrene)-3'-*O*-(2'-cyanoethoxydiisopropylamino-phosphino)-5'-*O*-(4,4'-dimethoxytrityl)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine (4)

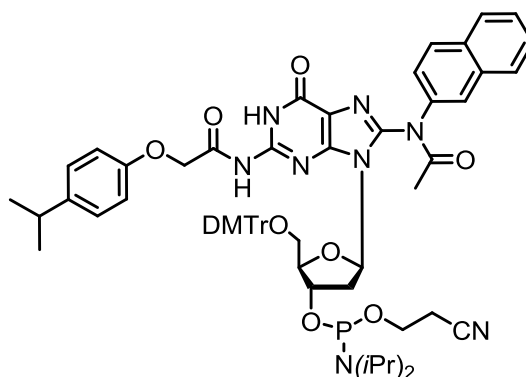


Yield: 99%

³¹P-NMR (162 MHz, CDCl₃): δ = 149.4, 148.8 ppm.

HR-MS (ESI⁻): m/z calculated for [C₆₉H₇₀N₈O₁₀P]⁻ ([M-H]⁻): 1201.4958, found: 1201.4951.

8-(*N*-Acetyl-2-aminonaphthalene)-3'-*O*-(2'-cyanoethoxydiisopropylamino-phosphino)-5'-*O*-(4,4'-dimethoxytrityl)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine (2)



Yield: 99%

³¹P-NMR (162 MHz, CDCl₃): δ = 149.4, 149.0 ppm.

HR-MS (ESI⁻): m/z calculated for [C₆₃H₆₈N₈O₁₀P]⁻ ([M-H]⁻): 1127.4801, found: 1127.4803.

References

- [1] L. C. Gillet, J. Alzeer, O. D. Scharer, *Nucleic Acids Res.* **2005**, 33, 1961-1969.
- [2] S. C. Koch, J. Kuper, K. L. Gasteiger, N. Simon, R. Strasser, D. Eisen, S. Geiger, S. Schneider, C. Kisker, T. Carell, *Proc. Natl. Acad. Sci. U. S. A.* **2015**, 112, 8272-8277.
- [3] W. Kabsch, *Acta Crystallogr. D Biol. Crystallogr.* **2010**, 66, 133-144.
- [4] a) K. Diederichs, P. A. Karplus, *Acta Crystallogr. D Biol. Crystallogr.* **2013**, 69, 1215-1222; b) P. Evans, *Science* **2012**, 336, 986-987; c) P. A. Karplus, K. Diederichs, *Science* **2012**, 336, 1030-1033.

- [5] G. N. Murshudov, P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 355-367.
- [6] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 486-501.
- [7] a) J. Painter, E. A. Merritt, *J. Appl. Crystallogr.* **2006**, *39*, 109-111; b) M. D. Winn, G. N. Murshudov, M. Z. Papiz, *Methods Enzymol.* **2003**, *374*, 300-321.
- [8] a) P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, et al., *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 213-221; b) P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart, P. D. Adams, *Acta Crystallogr. D Biol. Crystallogr.* **2012**, *68*, 352-367.
- [9] V. B. Chen, W. B. Arendall, 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 12-21.
- [10] E. Krissinel, K. Henrick, *J. Mol. Biol.* **2007**, *372*, 774-797.
- [11] R. Lavery, M. Moakher, J. H. Maddocks, D. Petkeviciute, K. Zakrzewska, *Nucleic Acids Res.* **2009**, *37*, 5917-5929.
- [12] E. Krissinel, K. Henrick, *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2256-2268.
- [13] M. Munzel, C. Szeibert, A. F. Glas, D. Globisch, T. Carell, *J. Am. Chem. Soc.* **2011**, *133*, 5186-5189.
- [14] a) L. C. Gillet, O. D. Scharer, *Org. Lett.* **2002**, *4*, 4205-4208; b) Z. Wang, C. J. Rizzo, *Org. Lett.* **2001**, *3*, 565-568.

3.3. Strukturelle Einblicke in die Erkennung von N^2 -Aryl- und C8-Aryl DNA-Schäden durch das Reparaturprotein XPA/Rad14

C. Ebert*, N. Simon*, S. Schneider, T. Carell, *ChemBioChem* **2017**, 18, 1379-1382. „*Structural insights into the recognition of N^2 -aryl- and C8-aryl DNA lesions by the repair protein XPA/Rad14*“. (* geteilte Erstautorenschaft)

3.3.1. Lizenz

Kopie der Publikation mit Erlaubnis des Verlags; Copyright 2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

3.3.2. Einführung

Der Erhalt der Kristallstrukturen von Rad14 mit verschiedenen C8-Acetylaryl geschädigten DNAs (Abschnitt 3.1. und 3.2.) war ein großer Meilenstein für die Erforschung der Funktion von XPA/Rad14 in der NER. Dennoch liegen keine strukturellen Informationen vor, wie XPA/Rad14 die nicht acetylierte Schäden und die N^2 -Addukte erkennt. Diese Publikation beschreibt die C8-AF-dG und die N^2 -AAN-dG Synthese, deren Einbau in die jeweiligen Oligonukleotide und die Kristallisation mit Rad14. Die erhaltenen Kristallstrukturen von Rad14 mit N^2 -AAN-dG und C8-AF-dG zeigen, dass die Schadenserkennung durch die Stabilisierung des DNA-Knicks durch XPA/Rad14 erfolgt.

3.3.3. Autorenbeitrag

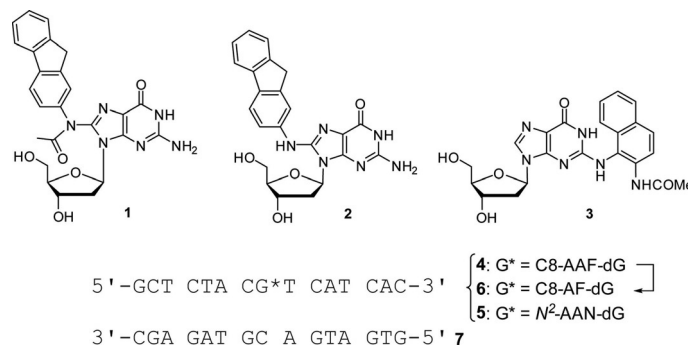
Für dieses Manuskript synthetisierte ich ein Oligonukleotid mit einem C8-AF-dG Schaden. Dafür führte ich die Synthese des C8-AAF-dG Schadens und dessen Einbau in ein Oligonukleotid durch. Anschließend führte ich die Deacetylierungsreaktion des C8-AAF-dG Schadens im Oligonukleotid durch, um den C8-AF-dG Schaden zu erhalten. *Charlotte Ebert* synthetisierte den N^2 -AAN-Schaden und dessen Oligonukleotid. Weiterhin reinigte ich das Rad14 und XPA Protein für die Kristallisations- und Bindungsstudien auf. Die Bindungsstudien, Schmelzpunktmessungen und Kristallisationsstudien wurden zusammen mit *Charlotte Ebert* durchgeführt. *Sabine Schneider* übernahm die Auswertung der Kristallstrukturdaten.

Structural Insights into the Recognition of N^2 -Aryl- and C8-Aryl DNA Lesions by the Repair Protein XPA/Rad14

Charlotte Ebert⁺,^[a] Nina Simon⁺,^[a] Sabine Schneider,^{*,[b]} and Thomas Carell^{*,[a]}

Aromatic amines are strongly carcinogenic. They are activated in the liver to give reactive nitrenium ions that react with nucleobases within the DNA duplex. The reaction occurs predominantly at the C8 position of the dG base, thereby giving C8-acetyl-aryl- or C8-aryl-dG adducts in an electrophilic aromatic substitution reaction. Alternatively, reaction with the exocyclic 2-NH₂ group is observed. Although the C8 adducts retain base-pairing properties, base pairing is strongly compromised in the case of the N^2 adducts. Here we show crystal structures of two DNA lesions, N^2 -acetylnaphthyl-dG and C8-fluorenyl-dG, within a DNA duplex recognized by the repair protein Rad14. The structures confirm that two molecules of the repair protein recognize the lesion and induce a 72 or 78° kink at the site of the damage. Importantly, the same overall kinked structure is induced by binding of the repair proteins, although the structurally different lesions result in distinct stacking interactions of the lesions within the duplex. The results suggest that the repair protein XPA/Rad14 is a sensor that recognizes flexibility. The protein converts the information that structurally different lesions are present in the duplex into a unifying sharply kinked recognition motif.

Aromatic amines are known to be strongly carcinogenic.^[1] The mode of action involves oxidation of the amines in the liver followed by their conversion into mono- or N,N -bis-acetyl compounds. After elimination of one acetyl group a reactive nitrenium ion compound is formed; this reacts with purine bases, mainly deoxyguanosine (dG), in DNA to give either C8 or N^2 bulky adducts (Scheme 1).^[1e,2] These bulky adducts are repaired by the nucleotide excision repair (NER) system.^[3] Lesion recognition is generally believed to proceed through binding of the NER protein XPC.^[4] Lesion binding was also reported for XPA, although it is assumed that XPA acts as a scaffold protein that is important for the assembly of the repair factors at the damaged site.^[3i,5] Patients with mutated XPA exhibit the strongest xeroderma pigmentosum phenotype, thus showing that the



Scheme 1. C8-AF-dG, N^2 -AAN-dG, and C8-AAF-dG lesions studied in complexes with Rad14; duplex used for crystallization (G* = lesion).

platform protein XPA is of paramount importance for efficient repair of bulky adduct lesions.^[6]

We recently crystallized yeast Rad14 (homologue of XPA) in complex with DNA containing C8-acetylaminofluorenyl bulky lesions (1).^[7] Because the 3D structures of XPA and Rad14 are very similar, structural insights gained with Rad14 can be directly related to lesion recognition in the human NER system.^[5d,e] The Rad14 structures clarified that Rad14 binds (2:1) to the lesion-containing DNA and that upon binding a sharp DNA kink of about 70° is induced. Because the NER process requires the assembly of a large number of proteins at the damage site to excise a stretch of 20–24 nucleobases around the lesion,^[3i,j,8] we speculated that the kink is a major recognition signal that enables recruitment of the proteins involved in the repair reaction.

Although structures of various acetylated C8-acetylaryl-dG lesions such as 1 have been solved,^[7] the question remains as to how XPA/Rad14 recognizes other nonacetylated lesions, like 2 and particularly N^2 bulky adducts such as 3 (Scheme 1). Here we obtained crystal structures of the DNA-binding domain of Rad14 in complex with DNA duplexes containing the non-acetylated C8-aminofluorene lesion (2, C8-AF-dG) and the N^2 -acetylnaphthyl adduct (3, N^2 -AAN-dG). We show that the ultimate recognition mode (2:1 binding) and the induction of a large kink at the lesioned site is preserved, despite distinct stacking interactions of the lesions in the duplex within the (Rad14)₂:DNA complex.

First we synthesized phosphoramidites of the C8-AAF-dG and of the N^2 -AAN-dG adducts according to published procedures.^[7b,9] We used phosphoramidite-based solid phase DNA synthesis to prepare oligonucleotides 4 and 5 (Scheme 1).^[10] Oligonucleotide 6 was obtained by deacetylation of 4.^[11] 5 and 6 were subsequently annealed with the complementary strand 7 to give duplexes. These were crystallized with a truncated

[a] C. Ebert,⁺ N. Simon,⁺ Prof. Dr. T. Carell
Center for Integrated Protein Science at the Department of Chemistry
Ludwig-Maximilians-Universität München
Butenandtstrasse 5–13, 81377 München (Germany)
E-mail: thomas.carell@lmu.de

[b] Dr. S. Schneider
Center for Integrated Protein Science at the Department of Chemistry
Technische Universität München
Lichtenbergstrasse 4, 85748 Garching (Germany)
E-mail: sabine.schneider@tum.de

[⁺] These authors contributed equally to this work.

Supporting information and the ORCID identification numbers for the authors of this article can be found under <https://doi.org/10.1002/cbic.201700169>.

Rad14 (Rad14_{188–302}, the minimal DNA-binding domain).^[7a] In 2014, Sugitani and others reported that DNA-binding strength is increased if the binding domain is extended at the C terminus by 20 amino acids (for XPA).^[12] Although this extended version gives better DNA binding, our shorter constructs crystallized well and afforded insight into the binding mode. We noticed, however, that in comparison to crystals with oligonucleotides containing lesion 1,^[7a] the crystals for this study grew significantly more slowly. In the case of Rad14_{188–302} and duplex 5:7, we had to wait three months before crystals were obtained, and two months for duplex 6:7, thus indicating that the system needs time to find the thermodynamically most favorable situation. We noted no protein or DNA degradation during this time (Supporting Information).

We first prepared 15-mer duplexes with C8-AF-dG and N²-AAN-dG (Table S1 in the Supporting Information) and confirmed binding with an electrophoretic mobility shift assay (EMSA). Lesion-containing duplexes were nicely recognized by XPA as expected (Figure 1A and B). Next we crystallized duplexes 5:7 and 6:7 with Rad14_(188–302) and obtained crystals that diffracted to 2.2 Å (5:7, C8-AF-dG lesion; in 0.22 M ammonium nitrate and 38% MPD) and to 1.9 Å (6:7, N²-AAN-dG lesion; in 0.02 M CaCl₂, 0.1 M NaOAc pH 4.6 and 30% MPD). As observed in all Rad14–DNA complex structures determined so far, the DNA bound (1:1) in two orientations, with a pseudo-C₂ symmetry axis running directly through the kink (Figure 1C and D). Thus, the lesion-containing strand is on one side of the duplex in 50% of cases, and on the other side in the other

50%. For clarity, we show only one DNA duplex with the lesion located on one side.

Figure 1 shows the complexes between Rad14 and DNA containing C8-AF-dG (Figure 1C) or N²-AAN-dG (Figure 1D). They are, as expected, very similar. In both cases two Rad14 molecules binding to the lesion (one above, one below), and this induces a sharp kink of the duplex at the lesion site (> 70°). The DNA is kinked into the major groove, and the protein makes no direct interaction with the lesion. Despite the different lesions, the sharply kinked DNA motif induced by binding of two Rad14 molecules is similar, and we postulate that this motif induces the assembly of the full set of repair enzymes at the damaged site. This could explain why XPA/Rad14 is able to function as a scaffold protein in NER and how lesions with quite different structures induce the same NER process.

The fact that both damaged DNAs lead to the same overall structure is quite interesting if we compare the chemical structures of the C8-AF and N²-AAN lesions. In principle, C8-AF can still base pair like dG; this is however impossible for the N²-AAN-dG lesion, in which the aromatic unit fully blocks the Watson–Crick base pairing side.

In the case of the C8-AF-dG structure (Figures 2 and S1), the lesion adopts a *syn* conformation around the nucleosidic bond. The Watson–Crick face of the dG base points into the major groove, which is strongly compressed due to the 78° kink into this groove. The dG part of the lesion stacks effectively with the dT base on the 3' side (3.2 Å), whereas stacking with the 5' dC is completely abolished. The fluorene ring system is not in plane with the dG part: the two planes are oriented at 50°. The fluorene ring points out into the minor groove, which is strongly widened due to the sharp kink. Stacking interactions are established with the dA base of the complementary strand (3.9 Å). There is no stacking interaction with the dC:dG base pair below. In fact the fluorene unit stands basically in an edge-to-face orientation on this base pair (Figure 2B). This dC:dG base pair is consequently not flat but buckled (V-type orientation, an angle of 35°). But, the dC:dG base pair is clearly formed with three H-bonds (Figure 2A); this is not the case for the dT:dA base pair above (Figure 2B). Although the dA base is in place with the Watson–Crick side pointing towards the helical axis, the dT base 3' to the lesion is shifted out of the helix, almost flipped out into the compressed major groove. This maximizes the stacking interactions with the dG part of the lesion. Again, in the duplex, optimizing stacking interactions is energetically more important than establishing H-bonding to the complementary base. The location of the dC opposite the lesion is noteworthy. Instead of pairing, the lesion adopts the described *syn* conformation around the nucleosidic bond. This places the fluorene moiety towards the dC complementary base, which is consequently flipped out of the duplex. Optimizing stacking of the large π surface of the fluorene ring system is clearly more important than stacking of the small pyrimidine dC.

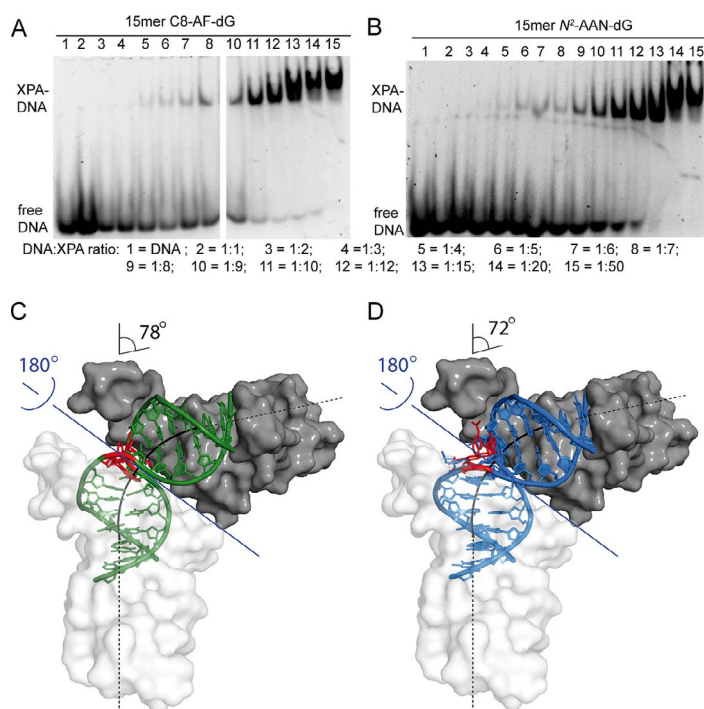


Figure 1. EMSA of XPA with DNA containing A) C8-AF-dG and B) N²-AAN-dG. Structures of the complexes between two Rad14 repair proteins (DNA-binding domains) bound to DNA containing C) C8-AF-dG or D) N²-AAN-dG. In both structures the DNA is kinked (78 and 72°, respectively). The pseudo-C₂-symmetry axis is shown in blue; the lesion (red) is positioned at the outer edge of the kink.

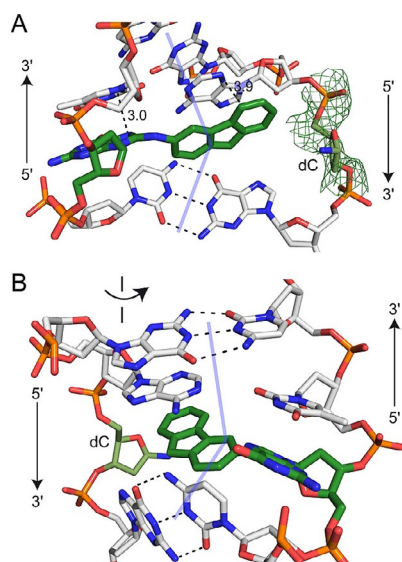


Figure 2. Detail of the C8-AF-dG crystal structure, with DNA as a stick model and the lesion highlighted in green. A) The C8-AF-dG lesion adopts a *syn* conformation around the nucleosidic bond and stacks with the dA of the complementary strand. The dG base stacks with the 3' dT base (the 3' dT:dA base pair is not formed). For the dC complementary base only the phosphodiester backbone and ribose moiety are defined in the $F_o - DF_c$ electron density (green mesh, contoured at 2.5σ). B) The 5' dC:dG base pair is formed despite buckled distortion. The direction of the kink is indicated in light blue.

We were unable to detect any electron density for the cytosine heterocycle of the dC base, thus suggesting that the heterocycle does not have a fixed conformation (Figure 2A).

The structure of the N^2 -AAN-dG differs at the lesion site. Although the lesion looks much more helix-disrupting due to the blocked Watson-Crick face, the crystal structure shows that the disrupting effect is much less than for the C8-AF-dG lesion. First, and most importantly, both the surrounding dC:dG and the dT:dA base pairs are intact (Figure 3A). These two base pairs are directly at the kink site, so they have a buckled rather than parallel orientation. The lesion slides into the V (formed by the buckling) to form stacking interactions with both base pairs thereby filling the void volume. This is possible because the dG part of the lesion and the naphthyl part are almost in a plane, so one large π surface is created (the angle between the two planes is only 10°). Stacking of the lesion is optimal with the dT:dA base pair on top, the dG packing below the dT base, and the naphthyl part stacking below the dA (Figure 3B). In this situation the naphthyl part occupies the space that is reserved usually for the dC complementary base, which is flipped out of the duplex into the much widened minor groove (Figure 3A). Due to the extensive stacking interactions of the N^2 adduct, the melting temperature of the DNA duplex is only marginally reduced compared to the undamaged control strand and the C8 adducts (Table S3).

The *N*-acetyl subunit of the lesion points in the same direction out of the minor groove. Again, electron density for the dC-heterocycle was not present in the structure, thus showing that the heterocycle is flexible. In addition to the flipped-out complementary dC base, the kink at the lesion site is another shared feature (this was already observed for the C8-AAF-dG

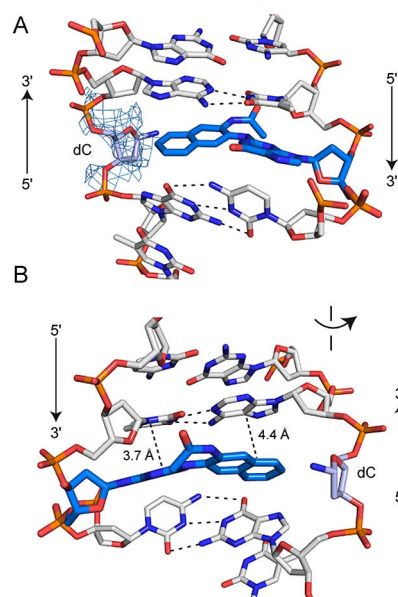


Figure 3. Detail of the N^2 -AAN-dG crystal structure, with DNA as a stick model, and the lesion highlighted in blue. A) The N^2 -AAN-dG lesion adopts an *anti* conformation around the nucleosidic bond. The base pairs above and below the lesion are intact. The dG base and the N^2 -AAN lesion are almost planar. For the dC complementary base only the phosphodiester backbone and ribose moiety are defined in the $F_o - DF_c$ electron density (blue mesh, contoured at 2.5σ). B) The lesion stacks with the dA of the complementary strand; the dG part stacks with the dT base above.

base), and hence it is possible that this dC is another structural element that induces assembly of the other repair factors at bulky adduct damaged sites. It is possible that this flipped out base functions as the recognition motif that allows the NER endonucleases to cleave specifically the damaged strand. Currently, we do not understand how the NER system distinguishes the damaged strand for incision from the undamaged strand. The Rad14:DNA structures are pseudo- C_2 symmetric, with the DNA bound 50:50 in both directions (Figure 1C and D). It is therefore likely that the flipped-out dC allows downstream factors to differentiate the damage-containing strand from the intact strand. This hypothesis needs investigation.

A comparison between the structures obtained here (C8-AF-dG lesion) and our previous work (C8-AAF-dG) offers an interesting perspective (Figure 4). The C8-AAF-dG lesion also exists in a *syn* conformation, but, whereas the two ring systems of the C8-AF-dG deviate from planarity by 50° , the two rings of the C8-AAF-dG are almost in a plane (perfect π surface). This allows the C8-AAF-dG lesion to form much better stacking interactions, not only with the upper dT:dA base pair but also with the lower dC:dG. C8-AAF-dG therefore improves the stability of the DNA kink significantly, and this might explain why this lesion is much better recognized by XPA/Rad14. In fact, for our XPA construct we noticed that when we performed the EMSA binding studies, the difference between binding to DNA containing lesions **2** and **3** and undamaged DNA is small (Figure S2), whereas for the C8-AAF-dG lesion, duplexes containing this lesion are more strongly bound.^[7a] Our crystal structures therefore confirm that the flexibility associated with the presence of a damaged base, as needed to establish a sharp 70–

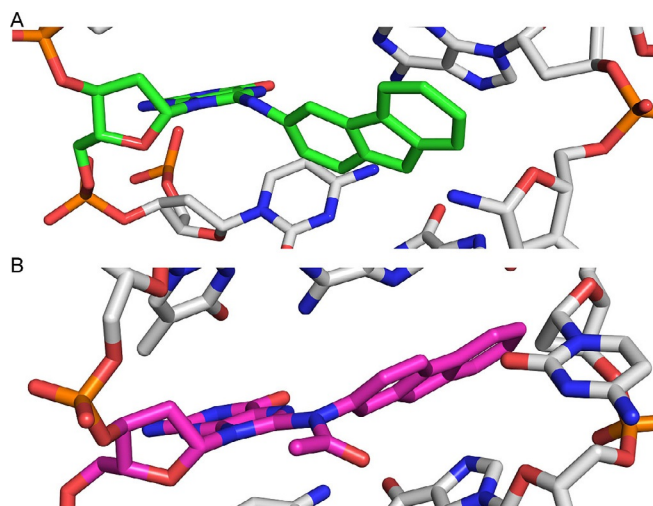


Figure 4. Detail of the A) C8-AF-dG (present study) and B) C8-AAF-dG (PDB ID: 5A3D) crystal structures. The lesions are highlighted in green and magenta, respectively. Both lesions adopt a *syn* conformation around the nucleosidic bond. The dG base and the C8-AAF lesion are almost planar; the C8-AF lesion and dG are in planes oriented at 50°.

78° kink, is a major factor for efficient recognition by XPA/Rad14. XPA/Rad14 is, in this sense, a duplex flexibility sensor.

The crystal structures have been deposited in the PDB under IDs: 5LCL and 5LCM.

Acknowledgements

The authors would like to thank the crystallization facility at the Max Planck Institute of Biochemistry for setting up the crystallization plates. We also want to thank the Swiss Light Source (SLS) and European Synchrotron Radiation Facility (ESRF) for beam time and the staff of beamlines PXI (SLS), MASSIF-3 and ID30B (ESRF) for assistance with the crystal testing and data collection. We thank the Deutsche Forschungsgemeinschaft, SFB1032 (TP-A5), SFB749 (TP-A4), SPP1784 and the Excellence Cluster CiPS^M for financial support and Dr. Markus Müller for critical reading of the manuscript and helpful advice.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: bulky adducts • DNA damage • nucleosides • nucleotide excision repair • Rad14 • XPA

- [1] a) J. A. Miller, *Cancer Res.* **1970**, *30*, 559–576; b) J. A. Miller, E. C. Miller, *Prog. Exp. Tumor Res.* **1969**, *11*, 273–301; c) R. H. Wilson, F. DeEds, A. J. Cox, Jr, *Cancer Res.* **1941**, *1*, 595–608; d) P. C. Howard, D. A. Casciano, F. A. Beland, J. G. Shaddock, Jr, *Carcinogenesis* **1981**, *2*, 97–102; e) R. H. Heflich, R. E. Neft, *Mutat. Res.* **1994**, *318*, 73–114; f) J. L. Radomski, E.

Brill, *Science* **1970**, *167*, 992–993; g) J. H. Weisburger, E. K. Weisburger, *Pharmacol. Rev.* **1973**, *25*, 1–66.

- [2] a) D. M. Grant, N. C. Hughes, S. A. Janezic, G. H. Goodfellow, H. J. Chen, A. Gaedigk, V. L. Yu, R. Grewal, *Mutat. Res. Fund. Mol. M.* **1997**, *376*, 61–70; b) G. R. Hoffmann, R. P. P. Fuchs, *Chem. Res. Toxicol.* **1997**, *10*, 347–359; c) J. A. Miller, E. C. Miller, *Environ. Health Perspect.* **1983**, *49*, 3–12; d) L. Möller, *Environ. Health Perspect.* **1994**, *102*, 139–146; e) D. Kim, F. P. Guengerich, *Annu. Rev. Pharmacol.* **2005**, *45*, 27–49; f) R. A. McClelland, A. Ahmad, A. P. Dicks, V. E. Licence, *J. Am. Chem. Soc.* **1999**, *121*, 3303–3310; g) F. A. Beland, W. T. Allaben, F. E. Evans, *Cancer Res.* **1980**, *40*, 834–840; h) W. G. Humphreys, F. F. Kadlubar, F. P. Guengerich, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 8278–8282; i) A. Dipple, *Carcinogenesis* **1995**, *16*, 437–441; j) S. A. Kennedy, M. Novak, B. A. Kolb, *J. Am. Chem. Soc.* **1997**, *119*, 7654–7664.
- [3] a) T. Lindahl, P. Karran, R. D. Wood, *Curr. Opin. Genet. Dev.* **1997**, *7*, 158–169; b) J. H. J. Hoeijmakers, *Nature* **2001**, *411*, 366–374; c) O. D. Schärer, *Angew. Chem. Int. Ed.* **2003**, *42*, 2946–2974; *Angew. Chem.* **2003**, *115*, 3052–3082; d) N. E. Geacintov, S. Broyde, T. Buterin, H. Naegeli, M. Wu, S. Yan, D. J. Patel, *Biopolymers* **2002**, *65*, 202–210; e) D. Gunz, M. T. Hess, H. Naegeli, *J. Biol. Chem.* **1996**, *271*, 25089–25098; f) J.-C. Huang, D. S. Hsu, A. Kazantsev, A. Sancar, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 12213–12217; g) L. C. J. Gillet, O. D. Schärer, *Chem. Rev.* **2006**, *106*, 253–276; h) S. Lagerwerf, M. G. Vrouwe, R. M. Overmeer, M. I. Fouteri, L. H. F. Mullenders, *DNA Repair* **2011**, *10*, 743–750; i) O. D. Schärer, *Cold Spring Harbor Perspect. Biol.* **2013**, *5*, a012609; j) S. C. Koch, N. Simon, C. Ebert, T. Carell, *Q. Rev. Biophys.* **2016**, *49*, e5.
- [4] a) J.-H. Min, N. P. Pavletich, *Nature* **2007**, *449*, 570–575; b) T. Riedl, F. Hanaoka, J.-M. Egly, *EMBO J.* **2003**, *22*, 5293–5303; c) M. Volker, M. J. Moné, P. Karmakar, A. van Hoffen, W. Schul, W. Vermeulen, J. H. J. Hoeijmakers, R. van Driel, A. A. van Zeeland, L. H. F. Mullenders, *Mol. Cell* **2001**, *8*, 213–224; d) M. S. Luijsterburg, G. von Bornstaedt, A. M. Gourdin, A. Z. Politi, M. J. Moné, D. O. Warmerdam, J. Goedhart, W. Vermeulen, R. van Driel, T. Höfer, *J. Cell Biol.* **2010**, *189*, 445–463; e) K. Sugawara, J. M. Y. Ng, C. Masutani, S. Iwai, P. J. van der Spek, A. P. M. Eker, F. Hanaoka, D. Bootsma, J. H. J. Hoeijmakers, *Mol. Cell* **1998**, *2*, 223–232.
- [5] a) W. Yang, *Cell Res.* **2008**, *18*, 184–197; b) B. C. Feltes, D. Bonatto, *Mutat. Res. Rev. Mutat. Res.* **2015**, *763*, 306–320; c) C.-L. Li, F. M. Golebiowski, Y. Onishi, N. L. Samara, K. Sugawara, W. Yang, *Mol. Cell* **2015**, *59*, 1025–1034; d) T. Ikegami, I. Kuraoka, M. Saijo, N. Kodo, Y. Kyogoku, K. Morikawa, K. Tanaka, M. Shirakawa, *Nat. Struct. Mol. Biol.* **1998**, *5*, 701–706; e) G. W. Buchko, S. Ni, B. D. Thrall, M. A. Kennedy, *Nucleic Acids Res.* **1998**, *26*, 2779–2788.
- [6] a) M. Kaposi, *Ann. Dermatol. Venereol.* **1883**, *4*, 29–38; b) C. F. M. Menck, V. Munford, *Genet. Mol. Biol.* **2014**, *37*, 220–233.
- [7] a) S. C. Koch, J. Kuper, K. L. Gasteiger, N. Simon, R. Strasser, D. Eisen, S. Geiger, S. Schneider, C. Kisker, T. Carell, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 8272–8277; b) N. Simon, C. Ebert, S. Schneider, *Chem. Eur. J.* **2016**, *22*, 10782–10785.
- [8] M. G. Kemp, J. T. Reardon, L. A. Lindsey-Boltz, A. Sancar, *J. Biol. Chem.* **2012**, *287*, 22889–22899.
- [9] a) R. R. Bonala, M. C. Torres, S. Attaluri, C. R. Iden, F. Johnson, *Chem. Res. Toxicol.* **2005**, *18*, 457–465; b) L. C. J. Gillet, O. D. Schärer, *Org. Lett.* **2002**, *4*, 4205–4208; c) Z. Wang, C. J. Rizzo, *Org. Lett.* **2001**, *3*, 565–568.
- [10] L. C. J. Gillet, J. Alzeer, O. D. Schärer, *Nucleic Acids Res.* **2005**, *33*, 1961–1969.
- [11] S. Shibutani, R. Gentles, F. Johnson, A. P. Grollman, *Carcinogenesis* **1991**, *12*, 813–818.
- [12] a) N. Sugitani, S. M. Shell, S. E. Soss, W. J. Chazin, *J. Am. Chem. Soc.* **2014**, *136*, 10830–10833; b) B. Hilton, N. Shkriabai, P. R. Musich, M. Kvaratskhelia, S. Shell, Y. Zou, *Biosci. Rep.* **2014**, *34*, e00162.

Manuscript received: March 24, 2017

Accepted manuscript online: April 26, 2017

Version of record online: May 30, 2017

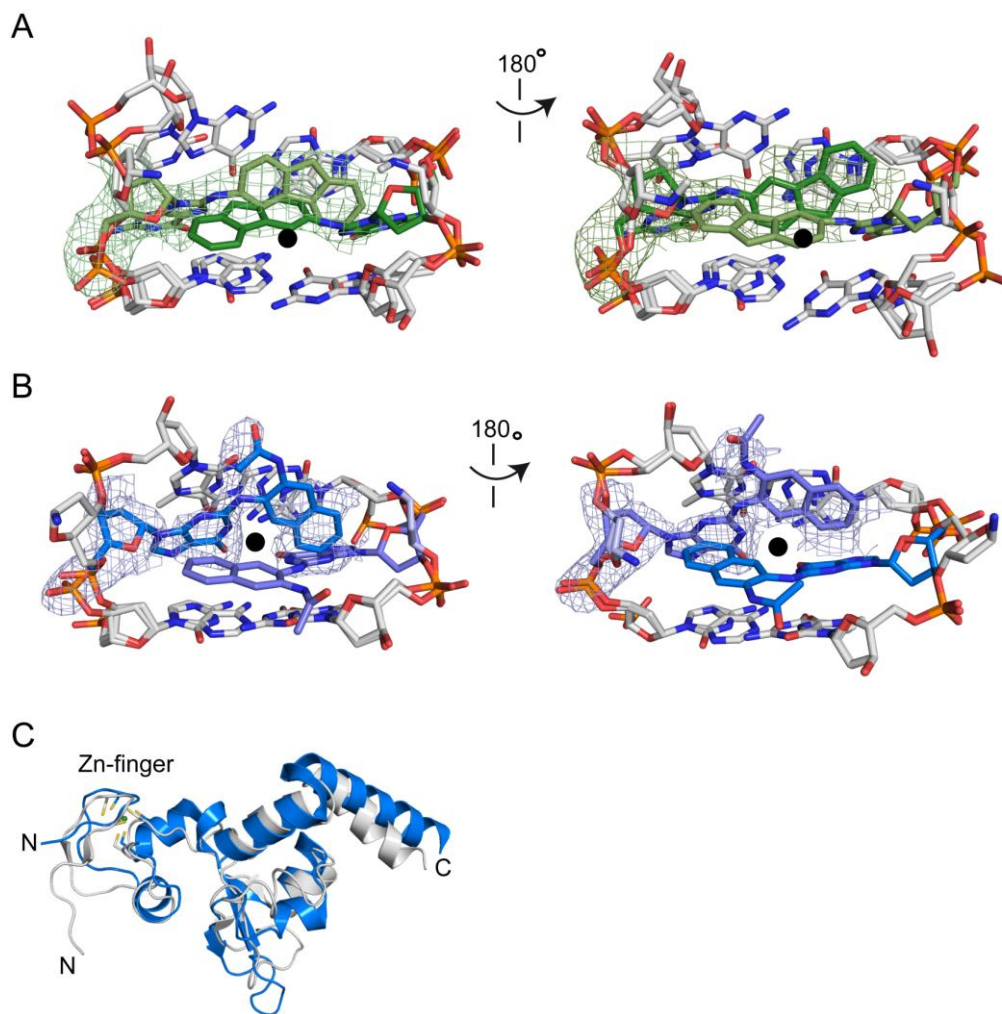
Supporting Information

Structural Insights into the Recognition of N^2 -Aryl- and C8-Aryl DNA Lesions by the Repair Protein XPA/Rad14

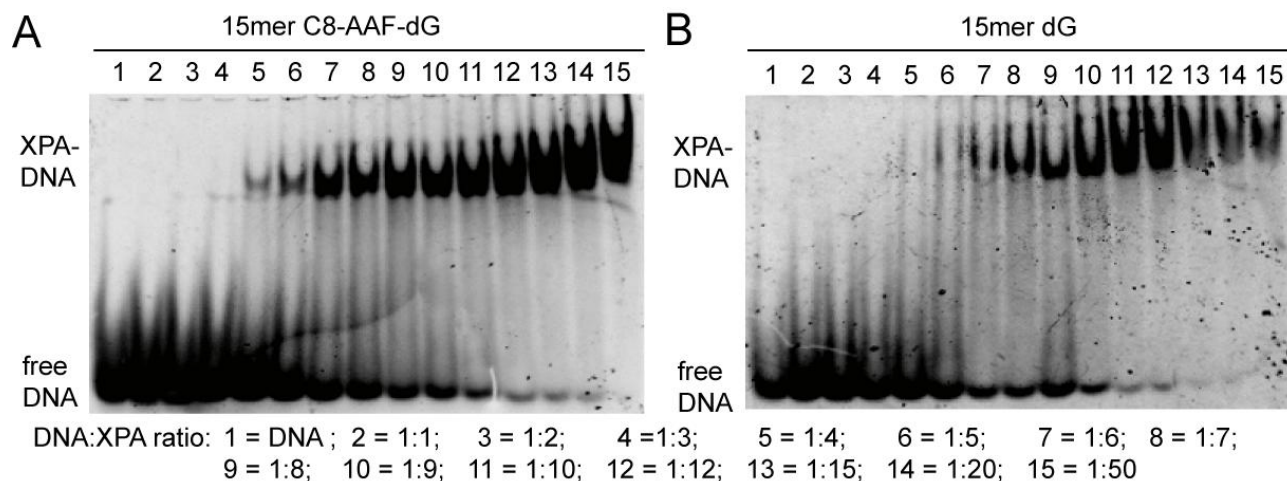
Charlotte Ebert^{+, [a]} Nina Simon^{+, [a]} Sabine Schneider,^{*, [b]} and Thomas Carell^{*, [a]}

cbic_201700169_sm_miscellaneous_information.pdf

Supporting Information



Supporting Figure S1. Simulated annealing-omit-difference F_o-DF_c electron density map of the C8-AF-dG (A) and N^2 -AAN-dG (B) lesions contoured at 2.5σ . The map was calculated after omitting the atoms of the lesions (both orientations). The two orientations of the DNA are overlaid in the figure and the black circle marks the C2-symmetry axis. C) Superposition of the DNA-binding domain of Rad14 (blue) and XPA (grey; NMR structure: PDB code 1XPA); r.m.s.d = 2.6 Å.



Supporting Figure S2. Electrophoretic mobility shift assay (EMSA) of XPA with (A) C8-AAF-dG and (B) dG containing DNA.

Supporting Table S1. Oligonucleotide sequences used in this study. The damaged guanine base is marked with a star and highlighted in bold.

ODN	Oligonucleotide sequence	G*
4	5'-GCT CTA CG *T CAT CAC-3'	C8-AAF-dG
5		N ² -AAN-dG
6		C8-AF-dG
7	5'-GTG ATG ACG TAG AGC-3'	-

Supporting Table S2. Data collection and refinement statistics. Statistics for the highest-resolution shell are shown in parentheses. * Indicates values calculated by the Molprobity server ^[1](<http://molprobity.biochem.duke.edu>)

	C8-AF-dG	N²-AAN-dG
Wavelength (Å)	0.968	0.976
Resolution range	49.3-2.2 (2.3 - 2.2)	49.2-1.9 (1.97- 1.9)
Space group	P 4 ₁	P 4 ₁
Unit cell	53.2 53.2 131.0	53.0 53.0 131.3
Total reflections	130,646 (9,621)	127,851 (12,486)
Unique reflections	18,314 (1,714)	28,333 (2,850)
Multiplicity	7.1 (5.6)	4.5 (4.4)
Completeness (%)	0.99 (0.93)	0.99 (0.99)
Mean I/sigma(I)	8.6 (1.6)	7.3 (1.4)
Wilson B-factor	31.4	28.4
R-merge	0.168 (0.901)	0.125 (0.704)
R-meas	0.181 (0.993)	0.142 (0.799)
R-pim	0.07 (0.408)	0.07 (0.373)
CC1/2	0.995 (0.566)	0.995 (0.604)
CC*	0.999 (0.85)	0.999 (0.868)
Reflections used in refinement	18,308 (1716)	28,349 (2,850)
Reflections used for R-free	895 (87)	1,370 (144)
R-work	0.216 (0.292)	0.231 (0.328)
R-free	0.257 (0.317)	0.2687 (0.331)
CC(work)	0.927 (0.72)	0.938 (0.719)
CC(free)	0.94 (0.68)	0.95 (0.65)
Number of non-hydrogen atoms		
Protein	1,877	1,632
DNA	1,157	1,156
Zinc-ions	2	2
Waters	49	112
Protein residues	227	229
RMS(bonds)	0.035	0.021
Bad bonds (protein)*	0/1,927 (0%)	0/1,967 (0%)
Bad bonds (DNA)*	3/1,217 (0.25%)	14/1,216 (1.15%)
RMS(angles)	2.1	2.5
Bad angles (protein)*	0/2598 (0%)	1/2,651 (0.04%)
Bad angles (DNA)*	33/1,861 (1.77%)	85/1,860 (4.57%)
Ramachandran favored (%)*	96	97
Ramachandran allowed (%)*	3.6	2.6
Ramachandran outliers (%)*	0	0.4
Rotamer outliers (%)*	2	1.4
Clashscore*	13.2	12.2
Average B-factor		
Protein	36.9	33.9
DNA	46.7	43.4
Zinc ions	24.4	23.0
Waters	30.2	33.2
Number of TLS groups	20	22

Supporting Table S3. Melting points.

DNA lesion	T _M [°C] 15mer
dG:dC	64
C8-AF-dG:dC	55
N ² -AAN-dG:dC	60
C8-AAF-dG:dC	48

Supporting Materials and Methods

Chemicals were purchased from Sigma-Aldrich, Fluka, Alfa Aesar, ChemGenes or Acros and used without further purification. Solutions were concentrated *in vacuo* on a Heidolph rotary evaporator. The solvents were of reagent grade and purified by distillation. Dry solvents were bought from Sigma-Aldrich. Acetonitrile for HPLC-purification was purchased from VWR. Water was purified by a Milli-Q Plus system from Merck Millipore. Chromatographic purification of products was accomplished using flash column chromatography on Merck Geduran Si 60 (40-63 µM) silica gel (normal phase). Thin layer chromatography (TLC) was performed on Merck 60 (silica gel F₂₅₄) plates. ¹H, ¹³C and ³¹P NMR spectra were recorded in deuterated solvents on Bruker ARX 300, Varian VXR400S, Varian Inova 400, Bruker AMX 600 and Bruker Avance 800 spectrometers and calibrated to the residual solvent peak. Multiplicities are abbreviated as follows: s = singlet, d = doublet, t = triplet, q = quartet, hept = septet, m = multiplet, br. = broad. For assignment of the structures, additional 2D NMR spectra (COSY, HSQC, HMBC) were measured. Matrix-assisted laser desorption/ionization-time-of-flight (MALDI-TOF) mass spectra were recorded on a Bruker Autoflex II. For MALDI-TOF measurements the samples were dialysed on a 0.025 µm VSWP filter (Millipore) against ddH₂O for 1 h and then cocrystallised in a 3-hydroxypicolinic acid matrix (HPA: 25 mg 3-hydroxypicolinic acid, 5 mg ammonium citrate in 0.5 mL H₂O). High resolution electrospray ionization mass spectra (HRMS-ESI) were recorded on a Thermo Finnigan LTQ-FT (ESI-FTICR). The pH-values of buffers were adjusted using a MP 220 pH-meter (Mettler Toledo). UV spectra and melting profiles were measured on a Jasco V-650 spectrometer using quartz glass cuvettes with 1 cm path length. The concentration of the purified DNA was determined with a NanoDrop ND-1000 spectrophotometer (PheqLab) from Thermo Scientific (1.5 µL sample volume) using calculated extinction coefficients from Oligo Calc (online: <http://www.basic.northwestern.edu/biotools/OligoCalc.html>) and applying Lambert-Beer's law. For strands containing artificial bases the extinction coefficient of their corresponding control strand was employed without corrections.

Phosphoramidite synthesis

The C8-AAF-dG phosphoramidite containing an isopropylphenoxyacetyl group at the N^2 -position for solid phase DNA synthesis and the N^2 -AAN-dG phosphoramidite were prepared as previously published.^[2]

DNA Preparation

The C8-lesion phosphoramidite containing an isopropylphenoxyacetyl group at the N^2 -position for solid phase DNA synthesis was incorporated into DNA using ultra mild conditions for DNA synthesis.^[3] AF-C8-dG lesion was prepared as previously published.^[4] The N^2 -AAN-dG phosphoramidite was incorporated into DNA using standard conditions.^[2c] Oligonucleotide synthesis was performed on an ABI 394 Nucleic Acid Synthesis System (Life Technologies, Darmstadt, Germany). Phosphoramidites for dA, dC, dG, dT and CPG carriers were obtained from Glen Research (Sterling, VA) or Link Technologies (Bellshill, Scotland). The coupling time for the modified phosphoramidites was extended to 3 x 15 min. The oligonucleotides were purified by reverse phase HPL chromatography (Macherey-Nagel, Nucleosil 100-7 C18, 10 × 250 mm, 0.1 M triethylammonium acetate, linear gradient from 0% to 40% in 45 min (C8-AAF-dG and C8-AF-dG) or 0% to 25% in 45 min (N^2 -AAN-dG). The purified strands were desalted using C18 cartridges (Waters, Sep-Pak C18 Classic Cartridge). The identity of the oligonucleotides were confirmed by matrix-assisted laser-desorption time-of-flight mass spectrometry and the purity analytical HPLC.

Electromobility shift assays

The damaged oligonucleotides (ODN **4-6**) were mixed with their complementary strands (ODN **7**) in equal molar amounts and annealed to form duplex DNA. The formation of the duplex DNA was checked by native TBE polyacrylamide gel electrophoresis. A standard DNA binding reaction mixture (10 μ L) contained 5 pmol of the C8- or N^2 -lesion duplex DNA and varying concentrations of purified XPAfl in binding buffer (50 mM HEPES-KOH pH 8.0, 100 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT and 5% (v/v) glycerol). After incubation at 37 °C for 30 min, 2 μ L loading dye (250 mM Tris-HCl pH 7.5, 40% glycerol, 0.2% bromphenolblue) were added and electrophoresis was performed on a 4% non-denaturing polyacrylamide TBE gel at 4 °C, which was pre-run for 30 min. Band shifts were visualized by incubating the gel in 1x TBE with SYBR Green I (1:10000, Sigma Aldrich) for 30 min and then detected via ImageReader.

Protein expression and purification

The *Saccharomyces cerevisiae* Rad14 DNA binding domain (residues 188-302) and the human XPA full length protein were expressed and purified as published in^[5]. The stability of the protein was tested via limited proteolysis^[5] and also proofed via gelfiltration.

Crystallization, data collection, structure determination and structure refinement

For co-crystallization experiments C8-AF and *N*²-AAN-dG containing 15mer DNA (ODN **4-6**, Supporting Table S1) was annealed in crystallization buffer to its counter strand ODN **7** (Metabion). Prior to crystallization protein and DNA were mixed in a molar ratio of 2:1 protein:DNA and incubated for 30 min at 4 °C. Crystals were grown in 0.22 M ammonium nitrate and 38 % 2-methyl-1,3-propanediol (MPD) (C8-AF-dG) or 0.02 M CaCl₂, 0.1 M NaOAc pH 4.6 and 30% MPD (*N*²-AAN-dG) at 4 °C, cryo-cooled in liquid nitrogen directly from the crystallisation drop and stored in liquid nitrogen until data collection.

Due to the small crystal size of 10-20 µm, diffraction data were collected with a micro beam at MASSIF-3 and ID30B beamlines, European Synchrotron Radiation Facility (ESRF)). Data processing was carried out with XDS^[6] and the resolution cut-offs were chosen according to the correlation coefficient of random half-data sets (1/2 CC) at about 50%.^[7] The coordinates of the Rad14-DNA complex (PDB code 5G34) were used to solve the structure by rigid-body refinement in REFMAC5^[8] after the atoms for the DNA were removed from the model. Completion of the models was done by iterative cycles of manual model building in COOT^[9] and restraint and TLS refinement^[10] in REFMAC5. The optimal number of TLS-groups was determined using the TLSMD-Server.^[10] Simulated-annealing omit electron density map coefficients were calculated using PHENIX.^[11] Structure validation and analysis was carried out with MOLPROBITY^[1], PISA^[12] and CURVES.^[13] For data processing and structure refinement statistics see the Supporting Table 2. Structural superposition were done with SSM^[14] and all structural figures prepared with PyMol (Delano Scientific).

Atomic coordinates and X-ray structure data were submitted to the Protein Data Bank (<http://www.ebi.ac.uk/pdbe/>). PDB codes: 5LCL (C8-AF-dG), 5LCM (*N*²-AAN-dG).

Melting Temperatures

The duplexes of ODN **4-6** and ODN **7** (Supporting Table S1, at concentration 2 µM) were dissolved in 250 µL of 10 mM Tris-HCl, pH 7.4, containing 150 mM NaCl. The samples were heated from 15 to 90 °C or cooled from 90 to 15 °C with a rate of 1 °C/min. Three melting profiles per sample were performed (Supporting Table S3).

References

- [1] V. B. Chen, W. B. Arendall, 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Cryst. D* **2010**, 66, 12-21.
- [2] a) Z. Wang, C. J. Rizzo, *Org. Lett.* **2001**, 3, 565-568; b) L. C. Gillet, O. D. Scharer, *Org. Lett.* **2002**, 4, 4205-4208; c) R. R. Bonala, M. C. Torres, S. Attaluri, C. R. Iden, F. Johnson, *Chem. Res. Toxicol.* **2005**, 18, 457-465; d) N. Simon, C. Ebert, S. Schneider, *Chem. Eur. J.* **2016**, 22, 10782-10785.
- [3] L. C. Gillet, J. Alzeer, O. D. Scharer, *Nucleic Acids Res.* **2005**, 33, 1961-1969.
- [4] S. Shibutani, R. Gentles, F. Johnson, A. P. Grollman, *Carcinogenesis* **1991**, 12, 813-818.
- [5] S. C. Koch, J. Kuper, K. L. Gasteiger, N. Simon, R. Strasser, D. Eisen, S. Geiger, S. Schneider, C. Kisker, T. Carell, *PNAS* **2015**, 112, 8272-8277.
- [6] W. Kabsch, *Acta Cryst. D* **2010**, 66, 133-144.
- [7] a) K. Diederichs, P. A. Karplus, *Acta Cryst. D* **2013**, 69, 1215-1222; b) P. Evans, *Science* **2012**, 336, 986-987; c) P. A. Karplus, K. Diederichs, *Science* **2012**, 336, 1030-1033.
- [8] G. N. Murshudov, P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Cryst. D* **2011**, 67, 355-367.
- [9] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Cryst. D* **2010**, 66, 486-501.
- [10] a) J. Painter, E. A. Merritt, *J. Appl. Crystallogr.* **2006**, 39, 109-111; b) M. D. Winn, G. N. Murshudov, M. Z. Papiz, *Methods Enzymol* **2003**, 374, 300-321.
- [11] a) P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Cryst. D* **2010**, 66, 213-221; b) P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart, P. D. Adams, *Acta Cryst. D* **2012**, 68, 352-367.
- [12] E. Krissinel, K. Henrick, *J Mol Biol* **2007**, 372, 774-797.
- [13] R. Lavery, M. Moakher, J. H. Maddocks, D. Petkeviciute, K. Zakrzewska, *Nucleic Acids Res.* **2009**, 37, 5917-5929.
- [14] E. Krissinel, K. Henrick, *Acta Cryst. D* **2004**, 60, 2256-2268.

3.4. Unveröffentlichte Arbeit: *In vivo* Studien zur Reparatur von AF-Schäden

3.4.1. Einleitung

Khobta et al. gelang es in *knock-out* Experimenten, die unterschiedlichen Reparaturraten von C8-AAF- und N^2 -AAF-dG zu erklären und zeigte dabei, dass XPA notwendig für die Reparatur beider Schäden ist.^[53] Die C8-AAF-Addukte können durch die globale Genomreparatur erkannt und entfernt werden, wobei das N^2 -AAF-Addukt nur durch die Transkriptions-gekoppelte Reparatur entfernt werden kann.^[53] Im Folgenden sollten diese Studien mit einem C8-AF-Addukt durchgeführt werden.

3.4.2. Projektbeitrag

Für dieses Projekt führte ich die Synthese des C8-AAF-dG Schadens durch. Die Oligonukleotidsynthese eines 18mers mit einem C8-AAF-dG Schaden wurde in Zusammenarbeit mit *Charlotte Ebert* durchgeführt. Anschließend führte ich die Deacetylierungsreaktion des C8-AAF-dG Schadens im Oligonukleotid durch, um den C8-AF-dG Schaden zu erhalten. Prof. Andriy Khobta und Nataliya Kitsera haben die *in vivo* Studien durchgeführt.

Die Ergebnisse dieser Studien sind bisher noch nicht veröffentlicht.

3.4.3. Material und Methoden

3.4.3.1. Synthese des 18mer Oligonukleotids mit einem C8-AF-dG Schaden

3.4.3.1.1. Synthese von C8-AAF-dG

Die Synthese des C8-AAF-dG Schadens erfolgte nach *Simon et al.* (s. Abschnitt 3.2. und 3.3.)^[142]

3.4.3.1.2. Oligonukleotidsynthese

Die DNA-Synthese zum Einbau der C8-AAF-dG Modifikation in einem 18mer (5'-CATTGCTTC[AAF-G]CTAGCACG-3') wurde an einem ABI 394 DNA/RNA Synthesizer (Applied Biosystems) durchgeführt, wobei die Standardprotokolle des Geräteherstellers verwendet wurden. Dabei wurden 1 μ mol oder

200 nmol Maßstäbe zur Herstellung verwendet. Für die DNA-Synthese wurden *ultramild*-Bedingungen nach *Gillet et al.*^[44] verwendet und die Kupplungszeit der Modifikation auf 30 min (2 x 15min) verlängert. Die Standard Phosphoramidite (Ac-dC, dT) und die *ultramild*-Phosphoramidite (Pac-dA, *i*Pr-Pac-dG), sowie die *ultramild* Q-CPG-Träger wurden von den Firmen Glen Research oder Link Technologies bezogen. Als Aktivator wurde der Aktivator 42 (Sigma Aldrich) und Acetonitril von der Firma Roth (Wassergehalt < 10 ppm oder < 30 ppm) verwendet. Die Konzentration der gelösten Phosphoramidite in Acetonitril betrug 0.1 M (< 10 ppm H₂O).

Tabelle 1: Verwendete Lösungen für die DNA-Synthese.

Lösungen	Zusammensetzung
Aktivator	Aktivator 42
<i>Capping</i> A	0.5 M Pac ₂ O in CH ₃ CN/2,6-Lutidin 9:1
<i>Capping</i> B	16% N-Methylimidazol in CH ₃ CN
Detritylierung	3% Dichloressigsäure (DCA) in DCM
Oxidation	2.5 mM I ₂ in 2,6-Lutidin in CH ₃ CN/H ₂ O 6:65:30 (v/v)

3.4.3.1.3. DNA-Entschützung

Durch die DMT-*off* Synthese wurde am Ende der DNA-Synthese die terminale 5'-DMTr Gruppe abgespalten. Zur Entschützung der Oligonukleotide und zur Abspaltung von der Festphase wurden *ultramild*-Bedingungen (Lösung aus 10% Diisopropylamin und 0.25 M β-Mercaptoethanol in Methanol, 55 °C, 18 h, 650 rpm) verwendet. Nach der Entschützung wurde die Lösung an einer SpeedVac eingengt.

3.4.3.1.4. Synthese von C8-AF-dG

Das C8-AF-dG Oligonukleotid (5'-CATTGCTTC[AF-G]CTAGCACG-3') wurde aus einem C8-AAF-dG Oligonukleotid nach *Shibutani et al.*^[143] hergestellt.

3.4.3.1.5. Aufreinigung des Oligonukleotids durch Hochleistungsflüssigkeitschromatographie (HPLC)

Zur Vorbereitung wurde das Oligonukleotid wieder in ddH₂O aufgenommen und durch einen 0.2 µm Spritzenfilter filtriert. Für die analytische *reversed phase* (RP) HPLC wurden die Geräte 2695 *Separations Module*, 2996 *Photodiode Array Detector* (Waters, Eschborn) sowie die RP-Säule EC 250/4 Nucleodur 100-3 C18 (Macherey-Nagel, Düren) verwendet. Weiterhin wurden für die präparative RP-HPLC die Geräte 2487 *Dual λ Absorbance Detector*, 1525 *Binary HPLC Pump* (Waters, Eschborn) und die RP-Säule VP 250/10 Nucleodur 100-5 C18 (Macherey-Nagel, Düren) verwendet. Zudem konnte die aromatische Aminofluoren Einheit bei einer zweiten Wellenlänge bei 305 nm detektiert werden.

Tabelle 2: Verwendete Puffer für die HPLC.

Lösungen	Zusammensetzung
Puffer A	0.1 M Triethylammoniumacetat (TEAA) in ddH ₂ O, pH 7.0
Puffer B	0.1 M TEAA in 80:20 CH ₃ CN: ddH ₂ O, pH 7.0

Zur Aufreinigung des 18mers wurde ein Gradient von 0–30% Puffer B in 45 min und einer Flussrate von 0.5 mL/min (analytische HPLC) oder 5 mL/min (präparative HPLC) verwendet. Das aufgereinigte Oligonukleotid wurde mittels Sep-Pak Kartuschen (Waters, Eschborn) nach dem Standardprotokoll des Herstellers entsalzt.

3.4.3.1.6. MALDI-TOF-Analyse des Oligonukleotids

Das entsalzte Oligonukleotid wurde auf einer HPA-Matrix (3-Hydroxypicolin-Säure (50 mg), 15-Krone-5 (10 µL), Ammonium-Hydrogencitrat (10 mg in 500 µL H₂O und 500 µL CH₃CN) ko-kristallisiert. MALDI-TOF Spektren wurden zur Identifizierung der richtigen Masse des Oligonukleotids mit einem Bruker-Autoflex II-Massenspektrometer und einem AnchorChip™-Target (Bruker MTP var/384) generiert.

3.4.3.1.7. Bestimmung der DNA-Konzentration

Die Konzentration des Oligonukleotids wurde mit dem NanoDrop Spektrophotometer ND-1000 (Peglab) bestimmt. Dabei wurde die Konzentration des gelösten Oligonukleotids durch die gemessene Absorption bei 260 nm und durch den Extinktionskoeffizienten bestimmt. Der Extinktionskoeffizient des Oligonukleotids wurde mit Hilfe des Lambert-Beerschen Gesetzes und aus der Summe der einzelnen Extinktionskoeffizienten berechnet (dA: $15.2 \text{ L mmol}^{-1} \text{ cm}^{-1}$, dC $7.1 \text{ L mmol}^{-1} \text{ cm}^{-1}$, dG $12.0 \text{ L mmol}^{-1} \text{ cm}^{-1}$, 8.4 $\text{L mmol}^{-1} \text{ cm}^{-1}$ und für C8-AF-dG wurde der Extinktionskoeffizient von dG angenommen).^[144]

3.4.3.2. In vivo Studien

Das C8-AF-dG 18mer Oligonukleotid wurde in den Vektor pZAJ-5c, der für EGFP kodiert, nach *Luhnsdorfer et al.* eingebaut.^[145] Die Überprüfung des erfolgreichen Einbaus vom Oligonukleotid wurde nach *Kitsera et al.* durchgeführt.^[53] Das Reporterkonstrukt mit dem DNA Addukt wurde in menschliche Wirtszellen transfiziert. Zudem wurde DsRed ko-exprimiert. Die EGFP Expression wurde 24 h nach der Transfektion mit einem Durchflusszytometrie-Gerät FACSCalibur™ (Beckton Dickinson, GmbH, Heidelberg) gemessen. Dabei wurde das Protokoll von *Khobta et al.* verwendet.^[146] Folgende Zelllinien wurden verwendet: XP-A (GM04312), XP-C (GM15983), CS-A (GM16094), CS-B (GM16095) und Reparatur-kompetente Zellen MRC-5 (AG10076) (alle von *NIGMS Human Genetic Cell Repository*, USA).

3.4.4. Ergebnisse und Diskussion

Die *in vivo* Studien zeigen die Reparatur von dG(C8)-AF in Zellen mit diversen NER-Defekten. Dabei sollte insbesondere die Relevanz von XPA für die Reparatur des Schadens untersucht werden. Die Zelllinie MRC-5 diene als Kontrolllinie, die EGFP zu 100% exprimiert. Die Ergebnisse der *in vivo* Studien des dG(C8)-AF Schadens zeigen, dass ein geringer Transkriptionsblock in XPA-defizienten Zellen vorliegt (Abb. 14). Dies bedeutet, dass 60% des Schadens von der RNA-Polymerase überlesen wird und somit EGFP exprimiert wird. Die verminderte Expression könnte daran liegen, das AF eine *syn*- oder *anti*-Konformation einnehmen kann.^[147] In der *syn*-Konformation kann es wie beim AAF-Schaden durch XPA erkannt werden und in der *anti*-Konformation kann der Schaden während der

Transkription überlesen werden. Für die Reparatur in der *syn*-Konformation ist XPA dann zwingend notwendig, wodurch es in den XPA-defizienten Zellen zum Transkriptionsblock von 40% kommt. Die XPC-defizienten Zellen sind für die Reparatur weniger relevant. Denn im Vergleich zu den XPA-defizienten Zellen, zeigen die XPC-defizienten Zellen, dass XPC für die Reparatur weniger relevant ist. Dies zeigt, dass XPA in der Schadenserkenkung eine wichtigere Rolle als XPC spielen könnte. Weiterhin ist erkennbar, dass eine Reparatur des Schadens in den CS-A und CS-B Zellen stattfindet, da diese ein ähnliches EGFP-Expressionslevel wie die Kontrollzelllinie aufweisen. Daher sind die CSA und CSB Proteine für die Reparatur des dG(C8)-AF Schadens nicht erforderlich, da dieser durch die globale Genomreparatur erkannt und entfernt wird.

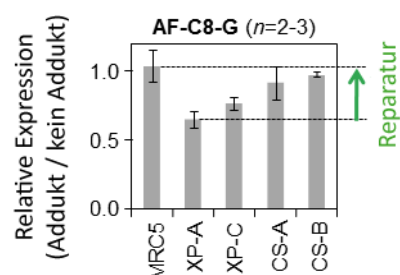


Abbildung 14: *In vivo* Studien vom C8-AF-dG Schaden. Dabei ist das Expressionslevel von EGFP in gesunden Zellen im Vergleich zu NER defekten Zellen dargestellt.

3.5. Unveröffentlichte Arbeit: Strukturelle Einblicke in die Erkennung von geschädigter DNA mit einer *Bubble* durch das Reparaturprotein XPA/Rad14

3.5.1. Einführung

Die Kristallstrukturen von Rad14 mit verschiedenen DNA Substraten (Abschnitt 3.1., 3.2. und 3.3.) zeigten erste funktionelle Einblicke von XPA/Rad14 in der NER. Dennoch liegen keine strukturellen Informationen vor, wie XPA/Rad14 geschädigte DNA mit einer *Bubble*, also dem Bereich, in dem der Doppelstrang bereits für die spätere Exzision aufgetrennt wurde, erkennt. Diese wäre von hohem Interesse, da eine *Bubble* in der DNA während des Präinziationskomplexes in der NER entsteht. Im Folgenden werden die Kristallisationsstudien von Rad14 mit unterschiedlich langen *Bubble*-DNA-Strukturen beschrieben. Dabei konnte eine Kristallstruktur von Rad14 mit einer 3nt *Bubble* in einem AAF-geschädigten 15mer erhalten werden. Diese zeigt, dass vor allem die Flexibilität der DNA ausschlaggebend für die Erkennung der Schäden durch Rad14 ist.

3.5.2. Projektbeitrag

Für dieses Manuskript synthetisierte ich den C8-AAF-dG Schaden und dessen Einbau in die Oligonukleotide. Weiterhin reinigte ich das Rad14 und XPA Protein für die Kristallisations- und Bindungsstudien auf. Auch führte ich die Bindungsstudien durch. Die Kristallisationsstudien wurden zusammen mit *Charlotte Ebert* angesetzt. *Sabine Schneider* übernahm die Auswertung der Kristallstrukturdaten.

3.5.3. Material und Methoden

3.5.3.1. Synthese der Oligonukleotide

Die Synthese des C8-AAF-dG Schadens erfolgte nach *Simon et al.* (s. Abschnitt 3.2. und 3.3.)^[142] Die Herstellung der AAF-geschädigten Oligonukleotide und deren Aufreinigung erfolgte wie in Abschnitte 3.5. beschrieben. Die ungeschädigten DNA-Stränge und Gegenstränge wurden von Metabion (Martinsried) bezogen. Tabelle 3 enthält eine Auflistung der einzelnen Sequenzen.

Tabelle 3: Auflistung der verwendeten Oligonukleotide für die Kristallisations- und Bindungsstudien. G*: C8-AAF-dG. Die rot markierten Basen zeigen die ungepaarten Fehlstellen, die bei der Hybridisierung mit dem Gegenstrang entstehen.

ODN	Bezeichnung	Nukleotidsequenz (5'→3')
1	15mer_UN	GCT CTA CGT CAT CAC
2	15mer_AAF	GCT CTA CG*T CAT CAC
3	15mer_Gegenstrang	GTG ATG TGC TAG AGC
4	25mer_UN	TCG ATG CAC AGA CAT CGC ACA CGT C
5	25mer_AAF	TCG ATG CAC AG*A CAT CGC ACA CGT C
6	25mer_Gegenstrang	GAC GTG TGC GAA CAG ACT GCA TCG A

3.5.3.2. Biochemische Studien

Die Reinigung der Proteine XPA und Rad14 erfolgte wie in Abschnitt 3.1., 3.2. und 3.3. beschrieben. Die Kristallisations- und Bindungsstudien wurden ebenso wie in Abschnitt 3.1., 3.2. und 3.3. durchgeführt. Für die Kristallisationsstudien wurde der MPD-Screen von Qiagen (Hilden) verwendet. Die Datenprozessierung und die Verfeinerungsstatistiken der Ko-Kristallstruktur befindet sich im Anhang unter 6.7.

3.5.3. Ergebnisse

Die Kristallstruktur von XPA/Rad14 sowohl mit einer *Bubble*-DNA als auch mit längerer DNA ist von hoher Relevanz, da diese weitere Einblicke in die Funktion von XPA liefern könnte. In diesem Abschnitt wurden Kristallisationsstudien mit Rad14, dem Hefehomolog von XPA, und zwei *Bubble*-DNAs unterschiedlicher Länge durchgeführt. Dafür konnte erfolgreich ein AAF-geschädigtes 15mer (**2**) und ein AAF-geschädigtes 25mer (**5**) synthetisiert werden. Dabei wurden die Gegenstränge so konzipiert, dass das ds 15mer ungepaarte Fehlstellen (**3**) enthält, welche zur Ausbildung einer 3nt *Bubble* führen, und dass das ds 25mer sechs ungepaarte Fehlstellen (**6**) enthält, welche zur Ausbildung einer 6nt *Bubble* führen. Nach Herstellung der ds Oligonukleotide, wurden diese sowie die ungeschädigten ds Oligonukleotide (**1** und **4**), in EMSA-Bindungsstudien mit dem Protein XPA verwendet (Abbildung 15). Dabei sollte die Bindungsaffinität von XPA an diesen Oligonukleotiden vor den Kristallisationsstudien bestätigt werden. In Abbildung 15 wird ersichtlich, dass XPA sowohl die ungeschädigte, als auch die AAF-geschädigte DNA bindet. Dabei bindet XPA an das 15mer bei einem Verhältnis von 2:1 (XPA:DNA) vermutlich schon als Dimer. Wohingegen XPA an das 25mer bei einem

Verhältnis von 1:1 zuerst als Monomer bindet und bei steigender Konzentration kooperativ ein zweites XPA bindet. Letztendlich bindet XPA sowohl die ungeschädigte als auch die AAF-geschädigte DNA im gleichen Verhältnis. Vermutlich erkennt und bindet XPA die Verzweigung der *Bubble*. Bereits *Yang et al.* zeigten, dass XPA stärker an verzweigte- (Y-förmige DNA) oder *Bubble*-DNA als an geschädigter DNA binden kann.^[148]

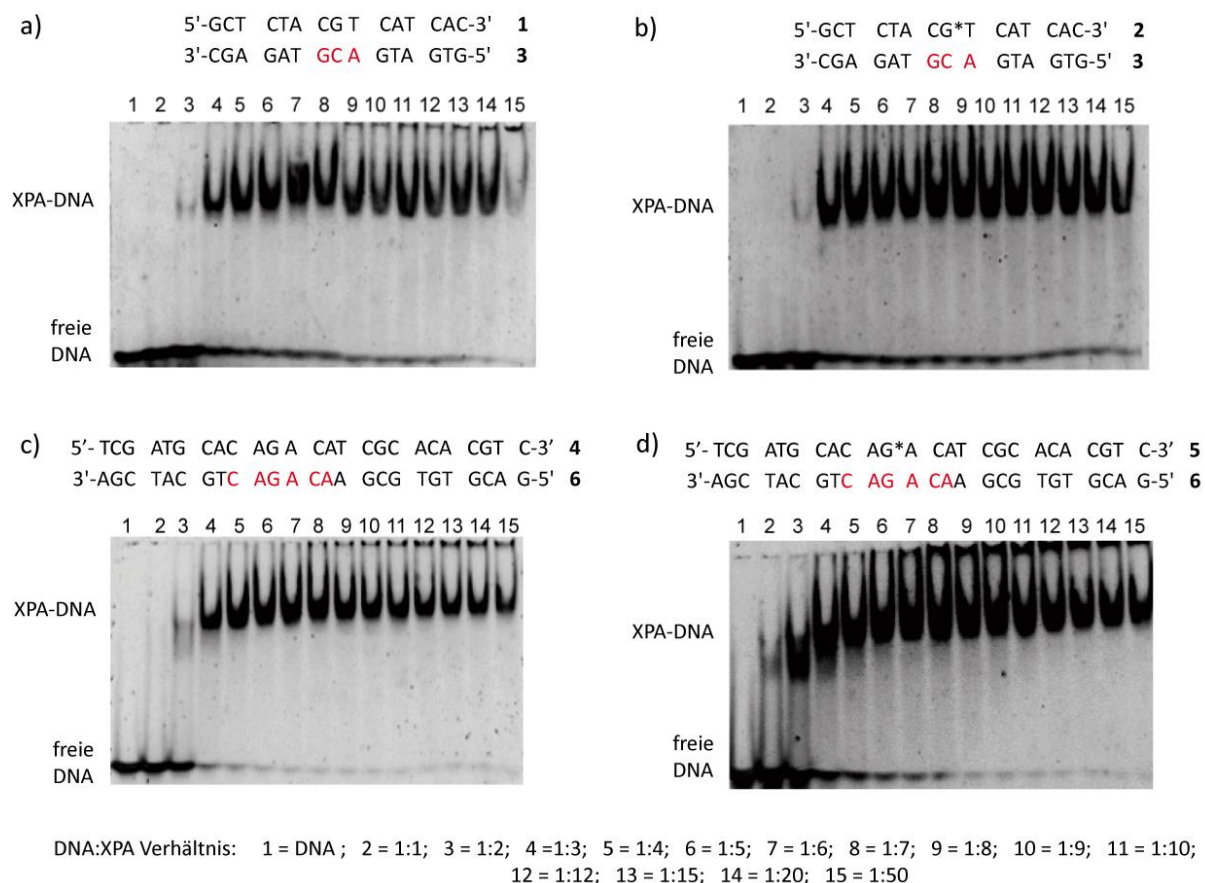


Abbildung 15: EMSA Gele der verschiedenen *Bubble*-DNAs. Die Basen in rot stellen die ungepaarten Basen dar, die zur *Bubble* Ausbildung führen. G*: C8-AAF-dG. Darstellung der Sequenz eines ungeschädigten (a) und eines AAF-geschädigten (b) 15mers mit 3 ungepaarte Fehlstellen, sowie das dazugehörige EMSA mit steigender Konzentration an XPA. Weiterhin die Darstellung der Sequenz eines ungeschädigten (c) und eines AAF-geschädigten (d) 25mers mit 6 ungepaarte Fehlstellen, sowie das dazugehörige EMSA mit steigender Konzentration an XPA.

In den Kristallisationsstudien mit Rad14 konnte für das AAF-geschädigte 15mer eine Kristallstruktur mit einer Auflösung von 2.4 Å (Abbildung 16) erhalten werden (Bedingungen: 2 M Ammoniumnitrat, 40% MPD und 4 °C). Diese ist identisch mit den zuvor erhaltenen Kristallstrukturen (Abschnitt 3.1., 3.2. und 3.3.). Auch hier binden zwei Rad14 Proteine an die DNA und diese ist um ca. 70 ° geknickt. Allerdings ist die Elektronendichte im Bereich des Schadens nicht gut aufgelöst, wodurch eine Zuordnung des Schadens in der Kristallstruktur nicht möglich ist. Auch die Positionen der ungepaarten Basen sind unzureichend definiert in der Kristallstruktur. Für eine bessere Zuordnung

könnten die ungepaarten Fehlstellen durch 5-Iodo-deoxyuridin ersetzt werden. Die schlechte Zuordnung des Schadens und der ungepaarten Basen in der *Bubble* könnten bedeuten, dass sich die π -stacking Interaktionen wahrscheinlich nicht ausbilden können und somit eine weniger relevante Rolle für die Schadenserkenkung spielen. Dadurch könnte vor allem die Flexibilität der DNA ausschlaggebend für die Erkennung der Schäden durch Rad14 sein.

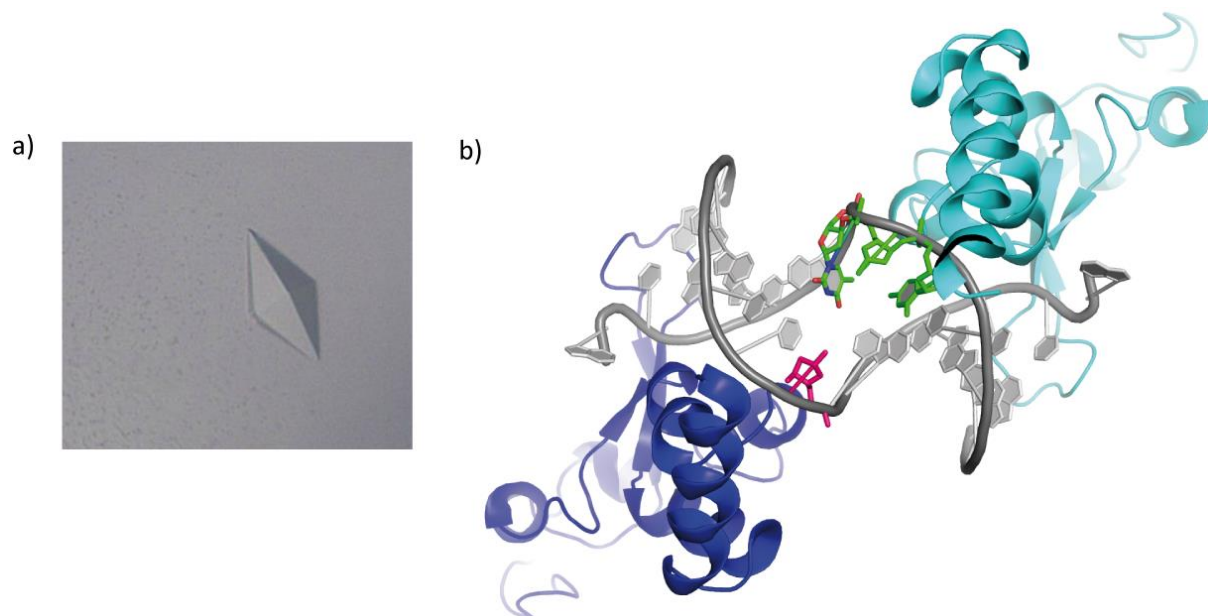


Abbildung 16: Ko-Kristallisation von $\Delta 4\text{CRad14s}$ mit einem AAF-geschädigten 15mer, welches 3 ungepaarte Fehlstellen zur *Bubble* Ausbildung besitzt. a) Erhaltene tetragonal bipyramidale Kristalle nach 3 Tagen bei 4 °C (2 M Ammoniumnitrat, 40% MPD und 4 °C). b) 2.5 Å Kristallstruktur von Rad14 mit dem 15mer. Die DNA ist in grau, der C8-AAF Schaden in pink und die ungepaarten Fehlstellen sind in grün markiert. Die DNA kann von zwei Rad14 Proteinen (in hell- und dunkelblau markiert) gebunden werden.

Auch in den Kristallisationsstudien mit Rad14 und dem AAF-geschädigten 25mer konnten Kristalle mit einer Auflösung von 2.5 Å erhalten werden (Bedingungen: 0.1 M Imidazol-HCl pH 8.0, 20% MPD und 4 °C). Hierbei konnte aber keine Struktur gelöst werden. Infolgedessen wurden die Kristalle aufgelöst und mittels Massenspektrometrie durch *Sabine Schneider* analysiert. Diese Analyse zeigte, dass Rad14 nicht in den Kristallen vorhanden war. Daneben konnten am Ende dieser Arbeit weitere Kristalle bei unterschiedlichen Bedingungen des MPD-Screens erhalten werden. Diese werden noch in Zukunft durch Röntgenstrukturanalyse noch analysiert werden.

4. Unveröffentlichte Arbeiten – Die humanen Proteine XPA und RPA

4.1. Einleitung und Projektbeiträge zu diesem Projekt

Einführung:

Das XPA Protein spielt eine wichtige Schlüsselrolle in der NER. Zudem kann es mit verschiedenen NER Proteinen, z.B. mit dem RPA Protein, interagieren. Bis heute wird die genaue Funktion und das Zusammenspiel von XPA und RPA in der NER untersucht. Vermutlich können sie unter anderem den NER Komplex zusammen stabilisieren oder auch eine Rolle bei der Schadenserkennung und bei der Schadensverifizierung spielen.^[124] Um diese Fragen beantworten zu können, wird im Nachfolgenden die Zusammensetzung eines XPA/RPA/DNA-Komplexes untersucht. Dabei sollte eine Methode entwickelt werden, die es erlaubt, den XPA/RPA/DNA-Komplex zu isolieren und verschiedene DNA-Substrate zu screenen. Die Struktur des isolierten Komplexes sollte dann durch Strukturanalysen gelöst werden, um weitere Einblicke in den NER-Mechanismus zu bekommen.

Projektbeitrag:

Für dieses Projekt synthetisierte ich den C8-AAF-dG Schaden. Ebenfalls stellte ich alle C8-AAF-dG geschädigten Oligonukleotide her. *Charlotte Ebert* unterstützte teilweise die Oligonukleotidsynthese. Alle weiteren biochemischen Versuche wurden zusammen mit *Johanna Bretzler* entwickelt und durchgeführt. Die Kristallisationsstudien wurden unter anderem in enger Zusammenarbeit mit *Prof. Caroline Kisker* und *Dr. Jochen Kuper* (Universität Würzburg) durchgeführt. Außerdem wurden die Kryo-Elektronenmikroskopie Studien in Kooperation mit der *Beckmann Gruppe* (LMU) und der *Dietz Gruppe* (TUM) durchgeführt. *Michael Stadlmeier* unterstützte die Arbeit mit *Crosslinking*-Studien. Um Zeit zu sparen, führte ich die Negativfärbung Experimente durch, während *Johanna Bretzler* unter Anleitung von *Sebastian Eustermann* die Evaluierung durchführte.

4.2. Material und Methoden

4.2.1. Material

4.2.1.1. Chemikalien und Lösungsmittel

Die Chemikalien und Lösungsmittel wurden von den Herstellern *AppliChem* (Darmstadt), *Duchefa Biochemie* (Haarlem, Niederlande), *Merck* (Darmstadt), *Roth* (Karlsruhe) oder *Sigma Aldrich* (Deisenhofen) bezogen. Alle Chemikalien für die biochemischen Studien waren in BioChemica Qualität.

4.2.1.2. Oligonukleotide

Die C8-AAF-dG geschädigten Oligonukleotide, die für biochemische Versuche verwendet wurden, sind in der folgenden Tabelle 4 aufgelistet. Die ungeschädigten Oligonukleotide sind ebenfalls aufgelistet und wurden von der Firma *Metabion* (Martinsried) bezogen.

Tabelle 4: Auflistung der ungeschädigten und C8-AAF-dG geschädigten Oligonukleotide. G*: C8-AAF-dG.

ODN	Bezeichnung	Nukleotidsequenz (5'→3')	Verwendung
15mer_1	15mer_AAF	GCTCTA CG*TCATCAC	Gelfiltrationsassay, EMSA, <i>Crosslinking</i>
15mer_2	15mer_UN	GCTCTA CGTCATCAC	Gelfiltrationsassay, EMSA, <i>Crosslinking</i>
15mer_3	15mer_GS	GTGATGACGTAGAGC	Gelfiltrationsassay, EMSA
15mer_4	15mer_GS_20nt_ov	GTCGTTCTGGAAGACCCACTGGTG ATGACGTAGAGC	Gelfiltrationsassay, EMSA, <i>Crosslinking</i>
25mer_1	25mer_AAF	TCGATGCACAG*ACATCGCACAC GTC	Gelfiltrationsassay, EMSA, Kristallisation, Kryo-EM
25mer_2	25mer_UN	TCGATGCACAGACATCGCACACG TC	Gelfiltrationsassay, EMSA, Kristallisation
25mer_3	25mer_GS	GACGTGTGCGATGTCTGTGCATC GA	Gelfiltrationsassay, EMSA
25mer_4	25mer_GS_20nt_ov	GTCGTTCTGGAAGACCCACTGGAC GTGTGCGATGTCTGTGCATCGA	Gelfiltrationsassay, EMSA, Kristallisation, Kryo-EM
25mer_5	25mer_GS_15nt_ov	TCGGAAGACCCACTGGACGTGTG CGATGTCTGTGCATCGA	Gelfiltrationsassay, EMSA

25mer_6	25mer_GS_10nt_ov	AGACCCACTGGACGTGTGCGATG TCTGTGCATCGA	Gelfiltrationsassay, EMSA
25mer_7	25mer_GS_6nt_2Y	GTCCACTGCGATGTCTGTGTGAG CT	Gelfiltrationsassay
33mer_1	33mer_AAF	CCAGGGGAGTGTGCGATG*CACA GATGTCCGAGCC	Gelfiltrationsassay, EMSA
33mer_2	33mer_UN	CCAGGGGAGTGTGCGATGCACAG ATGTCCGAGCC	Gelfiltrationsassay, EMSA
33mer_3	33mer_GS_13nt bubble	GGCTCGGACAAGACACGTAGCTG ACTCCCCTGG	Gelfiltrationsassay, EMSA
34mer_1	34mer_AAF	GTAGCTAGGTCGATG*CACAGAC ATCGCACACGTC	Gelfiltrationsassay, EMSA, Kristallisation, <i>Crosslinking</i> , Kryo-EM
34mer_2	34mer_UN	GTAGCTAGGTCGATGCACAGACA TCGCACACGTC	Gelfiltrationsassay, EMSA, <i>Crosslinking</i>
34mer_3	34mer_GS	GACGTGTGCGATGTCTGTGCATC GACCTAGCTAC	Gelfiltrationsassay, EMSA
34mer_4	34mer_GS_20nt_ov	GTCGTTCGGAAGAACCACTGGAC GTGTGCGATGTCTGTGCATCGAC CTAGCTAC	Gelfiltrationsassay, EMSA, Kristallisation, <i>Crosslinking</i> , Kryo-EM
34mer_5	34mer_GS_13nt_bubble	GACGTGTGCGAACAGACACGTA GCACCTAGCTAC	Gelfiltrationsassay, EMSA
34mer_6	34mer_GS_15nt_ov	TCGGAAGAACCACTGGACGTGTG CGATGTCTGTGCATCGACCTAGC TAC	Gelfiltrationsassay, EMSA
34mer_7	34mer_GS_10nt_ov	AGAACCACTGGACGTGTGCGATG TCTGTGCATCGACCTAGCTAC	Gelfiltrationsassay, EMSA
34mer_8	34mer_GS_1nt_ov	GGACGTGTGCGATGTCTGTGCAT CGACCTAGCTAC	Gelfiltrationsassay, EMSA
34mer_9	34mer_GS_1nt_ov	GACGTGTGCGATGTCTGTGCATC GACCTAGCTACG	Gelfiltrationsassay, EMSA
39mer_1	39mer_AAF	CCAGGGGAGTCACGTGCGATG*CA CAGAGTCTGTCCGAGCC	Gelfiltrationsassay, EMSA

39mer_2	39mer_UN	CCAGGGGAGTCACGTCGATGCAC AGAGTCTGTCCGAGCC	Gelfiltrationsassay, EMSA
39mer_3	39mer_GS_19nt_bubble	GGCTCGGACACTGAGACACGTAG CTGCACACTCCCCTGG	Gelfiltrationsassay, EMSA
45mer_1	45mer_UN	CCAGGGGAGTGTTACGTCGATG CACAGAGTCCATTGTCCGAGCC	Gelfiltrationsassay, EMSA
45mer_2	45mer_UN_25nt_bubble	GGCTCGGACATACCTGAGACACG TAGCTGCACTTGACTCCCCTGG	Gelfiltrationsassay, EMSA

4.2.1.3. Bakterien

Die, für die Klonierung, Mini-Plasmidpräparation und Expression von XPA bzw. RPA, verwendeten *E.coli* Stämme sind in Tabelle 5 aufgelistet. Der *E.coli* Stamm NEB5 α wurde bei der Amplifizierung von Plasmiden verwendet. Für die Expression von XPA wurde BL21(DE3)*pLysS* und für RPA wurde BL21(DE3) verwendet.

Tabelle 5: Verwendete *E.coli* Bakterienstämme.

<i>E. coli</i> Stamm	Genotyp	Firma
BL21(DE3)	$F^- ompT, hsdS_B (r_B^- m_B^-) gal dcm, (DE3)$	Novagen, Schwalbac
BL21(DE3) <i>pLysS</i>	$F^- ompT, hsdS_B (r_B^- m_B^-) gal dcm, (DE3), pLysS, (Cam^R)$	Novagen, Schwalbach
NEB5 α	<i>fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17</i>	NEB, Frankfurt

4.2.1.4. Enzyme

In dieser Arbeit verwendete Enzyme sind in Tabelle 6 aufgeführt:

Tabelle 6: Verwendete Enzyme.

Enzym	Verwendung	Firma
Benzonase®	Nuklease (250 U/ μ L), DNA-Verdau	Novagen, Schwalbach
<i>HindIII</i>	Restriktionsverdau	NEB, Frankfurt
Lysozym aus Hühnereiweiß	Zelllyse (100 mg/mL)	Sigma-Aldrich, Deisenhofen
<i>XbaI</i>	Restriktionsverdau	NEB, Frankfurt

4.2.1.5. Vektoren

Die Vektoren pENTRY-IBA10_XPAfl (bereitgestellt von *Sandra Koch*) und pPSG-IBA35 wurden aus der *StarGate®-Acceptor Vector Collection* (IBA, Göttingen) genutzt. Der Vektor p11d-tRPA wurde freundlicherweise von Prof. Dr. Marc Wold (Universität Iowa, USA) zur Verfügung gestellt. Diese Vektoren enthalten, neben verschiedenen Erkennungssequenzen für bestimmte Enzyme, auch Selektionsmarker als Resistenzgen wie z.B. Amp^R. Zur gezielten Aufreinigung besitzen die Proteine zur Epitopmarkierung (*tags*) ein N-terminales Peptid im pPSG-IBA35 Vektor. Die verwendeten Vektoren sind in Tabelle 7 dargestellt.

Tabelle 7: Verwendete Vektoren und Eigenschaften der Vektoren.

Bezeichnung	Größe	Resistenz	Affinitätstag	Verwendung
pENTRY-IBA10	1775 bp	Kan	-	Klonierung von <i>xpa</i>
pPSG-IBA35	2634 bp	Amp	N-terminaler 6xHis-tag	Expression von XPAfl
p11d-tRPA	9689 bp	Amp	-	Expression von RPA

In Abbildung 17 sind die Plasmidkarten von pPSG-IBA35-XPAfl und p11d-tRPA dargestellt.

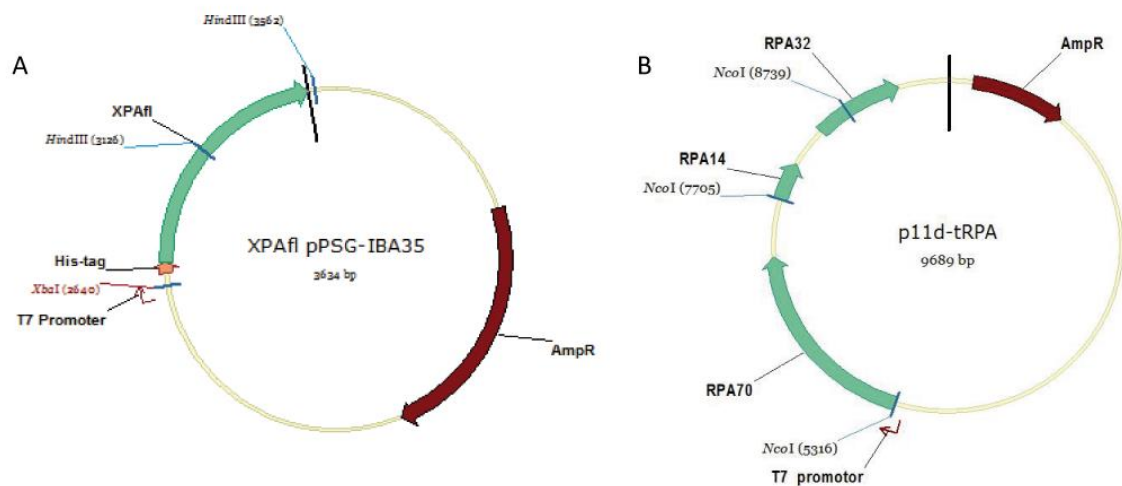


Abbildung 17: Vektorkarten von A) XPAfl pPSG-IBA35: AmpR: Ampicillin Resistenz; *HindIII* und *Xba*: Restriktionsstellen; N-terminaler 6xHis-Tag B) p11d-tRPA: AmpR: Ampicillin Resistenz; *NcoI*: Restriktionsstellen; drei Untereinheiten des Proteinkomplexes: RPA70, RPA32, RPA14.

4.2.1.6. Medien und Puffer

4.2.1.6.1. Medien zur Kultivierung von *E.coli*

LB-Medium	XPA:	Pepton aus Casein (10 g/L) Hefeextrakt (5 g/L) NaCl (7.5 g/L)
	RPA:	Pepton aus Casein (10 g/L) Hefeextrakt (5 g/L) NaCl (5 g/L) pH 7.5 (mit NaOH eingestellt)
Agar-Medium:		1.5% (w/v) Agar in LB-Medium 0.1% (v/v) Antibiotika-Stocklösung (nach autoklavieren)

4.2.1.6.2. Stocklösungen

Carbenicillin	100 mg/mL in ddH ₂ O (1000x Stock)
Chloramphenicol	34 mg/mL in EtOH (1000x Stock)
Kanamycin	50 mg/mL in ddH ₂ O (1000x Stock)
DTT	154 mg/mL in ddH ₂ O (1 M)
IPTG	238 mg/mL in ddH ₂ O (1 M)

4.2.1.6.3. Agarosegel

TAE-Laufpuffer:	40 mM Tris-Acetat pH 8.3
	2 mM Na ₂ -EDTA·2H ₂ O
TAE-Gelpuffer:	20 mM Tris pH 8.3
	1 mM Na ₂ -EDTA·2H ₂ O

5x DNA-Auftragspuffer: (in TAE gelöst)	50% (v/v) Glycerin 0.2% (w/v) SDS 0.05% (w/v) Bromphenolblau 0.05% (w/v) Xylencyanol
---	---

4.2.1.6.4. Native-PAGE

10x TBE-Puffer:	89 mM Tris-HCl pH 8.0 2 mM EDTA 89 mM H ₃ BO ₃
-----------------	--

EMSA-Auftragspuffer:	250 mM Tris-HCl pH 7.5 40% (v/v) Glycerin 0.2% (w/v) Bromphenolblau
----------------------	---

Bindungspuffer :	50 mM HEPES-NaOH pH 8.0 150 mM NaCl 5% (v/v) Glycerin 5 mM MgCl ₂ ·6H ₂ O 5 mM DTT 10 µM ZnCl ₂
------------------	---

EMSA-Laufpuffer:	0.25x TBE-Puffer
------------------	------------------

Detektionspuffer:	1x TBE-Puffer 0.01% (v/v) SYBR Green I
-------------------	---

4.2.1.6.5. SDS-PAGE

10x SDS-Laufpuffer:	0.25 M Tris 1.29 M Glycin 1% (w/v) SDS
---------------------	--

Sammelgelpuffer:	1.0 M Tris-HCl pH 6.8
Trenngelpuffer:	1.5 M Tris-HCl pH 8.8
5x SDS-Auftragspuffer:	200 mM Tris-HCl pH 6.8 8% (w/v) SDS 40% (w/v) Glycerin 0.04% (w/v) Bromphenolblau 400 mM DTT
APS-Lösung:	10% (w/v) Ammoniumperoxodisulfat
SDS-Lösung:	10% (w/v) SDS
Färbelösung I:	0.25% (w/v) Coomassie <i>Brilliant Blue</i> R 250 20% (v/v) AcOH 10% (v/v) EtOH
Entfärbelösung I:	20% (v/v) EtOH 10% (v/v) AcOH
Coomassie-Lösung II:	0.25% (w/v) Coomassie <i>Brilliant Blue</i> G 250 in EtOH
Färbelösung II:	5% (v/v) EtOH 7.5% (v/v) AcOH
Entfärbelösung II:	100% H ₂ O
DNA-Färbelösung:	1x TBE-Puffer 0.01% (v/v) SYBR Green I

4.2.1.6.6. XPA-Proteinreinigung mit einer Ni-NTA-Säule

Ni-NTA Puffer A:	25 mM Tris-HCl pH 7.5
	100 mM NaCl
	5 mM DTT
	10% (v/v) Glycerin
	10 μ M ZnCl ₂
	30 mM Imidazol

Ni-NTA Puffer B:	Ni-NTA Puffer A
	500 mM Imidazol

4.2.1.6.7. XPA-Proteinreinigung mit einer Heparin-Matrix

Heparin Puffer A:	25 mM Tris-HCl pH 7.5
	50 mM NaCl
	5 mM DTT
	10% (v/v) Glycerin
	10 μ M ZnCl ₂

Heparin Puffer B:	Heparin Puffer A
	1 M NaCl

4.2.1.6.8. XPA-Proteinreinigung mit einer Anionenaustaucher-Matrix (MonoQ)

MonoQ Puffer A:	25 mM Tris-HCl pH 7.5
	50 mM NaCl
	5 mM DTT
	10% (v/v) Glycerin
	10 μ M ZnCl ₂

MonoQ Puffer B:	MonoQ Puffer A
	1 M NaCl

4.2.1.6.9. XPA-Proteinreinigung mit Größenausschlusschromatographie (SEC)

Kristallisationspuffer II:	50 mM HEPES-NaOH pH 8.0
	150 mM NaCl
	5% (v/v) Glycerin
	5 mM MgCl ₂ ·6H ₂ O
	5 mM DTT
	10 µM ZnCl ₂

4.2.1.6.10. RPA-Proteinreinigung mit einer GEBLue-Matrix

2x HI Puffer:	60 mM HEPES-KOH pH 7.8
	0.50 % (w/v) Inositol
	0.02 % (v/v) NP-40
	0.5 mM EDTA
	2 mM DTT

HI Puffer A:	1x HI Puffer
--------------	--------------

HI Puffer B:	1x HI Puffer
	1 M KCl

HI Puffer C:	1x HI Puffer
	0.8 M KCl
	0.5 M NaSCN

HI Puffer D:	1x HI Puffer
	0.8 M KCl
	1.5 M NaSCN

4.2.1.6.11. RPA-Proteinreinigung mit einer HAP-Matrix

HI Puffer E:	1x HI Puffer 30 mM KCl
HI Puffer F:	1x HI Puffer 30 mM KCl 80 mM KH_2PO_4
HI Puffer G:	1x HI Puffer 30 mM KCl 500 mM KH_2PO_4
Lagerungspuffer 1 HAP:	10 mM KH_2PO_4 in H_2O
Lagerungspuffer 2 HAP:	10 mM KH_2PO_4 in EtOH

4.2.1.6.12. RPA-Proteinreinigung mit einer Anionenaustaucher-Matrix (MonoQ)

HI Puffer A:	1x HI Puffer
HI Puffer B:	1x HI Puffer 1 M KCl

4.2.1.6.13. Kristallisations-, Gelfiltrationsassay- und Kryo-EM-Puffer

Kristallisationspuffer I:	50 mM HEPES-NaOH pH 8.0 5% (v/v) Glycerin 50 mM $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$ 5 mM DTT 10 μM ZnCl_2
---------------------------	---

Kristallisationspuffer II:	50 mM HEPES-NaOH pH 8.0
	150 mM NaCl
	5% (v/v) Glycerin
	5 mM MgCl ₂ ·6H ₂ O
	5 mM DTT
	10 µM ZnCl ₂

4.2.1.6.14. Puffer zum Beladen einer neuen Ni-NTA-Säule

Stripping Puffer:	20 mM Na ₃ PO ₄ pH 7.4
	0.5 M NaCl
	50 mM EDTA

Binding Puffer:	20 mM Na ₂ PO ₄ pH 7.4
	0.5 M NaCl
	20 mM Imidazol

Nickel-Lösung:	0.1 M NiSO ₄
----------------	-------------------------

4.2.1.6.15. HPLC Puffer

Puffer A	0.1 M TEAA pH 7.0
----------	-------------------

Puffer B	0.1 M TEAA pH 7.0
	80% (v/v) CH ₃ CN

4.2.1.6.16. DNA-Synthese Lösungen

Aktivator	Aktivator 42
-----------	--------------

<i>Capping A</i> - Lösung	0.5 M Pac ₂ O
	90% (v/v) CH ₃ CN
	10% (v/v) 2,6-Lutidin

Capping B - Lösung

16% (v/v) *N*-Methylimidazol in CH₃CN

Detritylierung - Lösung

3% (v/v) DCA in DCM

Oxidation-Lösung

2.5 mM I₂

6% (v/v) 2,6-Lutidin

65% (v/v)

4.2.1.7. Protein- und DNA-Marker

In Abbildung 18 sind der verwendete Proteinstandard PageRuler™ *Unstained Protein Ladder* (Thermo Scientific, Dreieich) und die beiden genutzten DNA-Standards (NEB, Frankfurt) abgebildet.

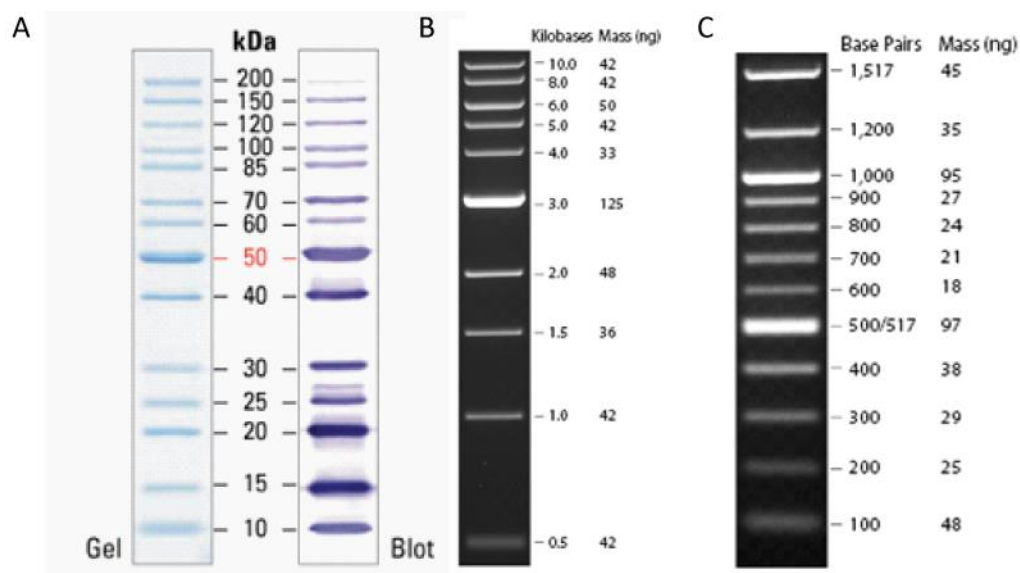


Abbildung 18: Verwendete Standards. A) Proteinstandard PageRuler™. B) DNA-Standard 1 kb. C) DNA-Standard 100 bp. ^[149-151]

4.2.1.8. Präparationskits

Zur Klonierung von *xpa* wurde das *StarGate*-Kit® (IBA, Göttingen) und zur Plasmidisolierung das *GeneJET* Plasmid Miniprep Kit (Thermo Scientific, Dreieich) verwendet.

4.2.1.9. Geräte

Die verwendeten Geräte sind in Tabelle 8 aufgeführt.

Tabelle 8: Verwendete Geräte in dieser Arbeit.

Geräte	Firma
ABI 394 DNA/RNA Synthesizer	Applied Biosystems, Waltham, USA
Agarose Gelelektrophoresekammer	Bio-Rad, München
ÄkTA™ Purifier Chromatographie System	GE Healthcare, München
Analysenwaage	Sartorius, Göttingen
Anaylsenwaage PG603-S	Mettler Toledo, Göttingen
Autoklav Laboklav ECO	HSP Steriltechnik, Detzel Schloss
Biofuge pico	Heraeus, Hanau
BioPhotometer 6131	Eppendorf, Hamburg
Eismaschine	Ziegra, Isernhagen
Image Reader LAS3000	Fujifilm, Tokio, Japan
HPLC	Waters, Eschborn
Inkubator 1S	Noctura, Wiesloch
Inkubator 44R	New Brunswick, Nürtingen
Lyophilie	Christ, Osterode am Harz
MALDI-TOF Autoflex II-Spektrometer	Bruker Daltonics, Bremen
Mastercycler Personal	Eppendorf, Hamburg
Mini Protean 4 Cell	Biorad, München
NanoDrop UV-Spektrometer	Peqlab, Erlangen
pH Meter MP220	Mettler Toledo, Gießen
Polymax Wave Motion Platform	Heidolph, Schwalbach
Power Supply	Biorad, München
Rotilabo-Mini Zentrifuge	Roth, Karlsruhe
Rotor BioSealF10S-4x 1000 LEX	Thermo Life Science, Dreieich
Rotor A-4-62	Eppendorf, Hamburg
Rotor FA-45-24-11	Eppendorf, Hamburg
Rotor Sorvall SS34	Thermo Life Science, Dreieich
Scanner	Epson, Japan
SDS Gelkammern	Biorad, München

Sonoplus HD 2070	Bandelin, Berlin
Sonorex RK 100 H	Bandelin, Berlin
Sorvall RC6 Plus Zentrifuge	Thermo Life Science, Dreieich
Speedvac	Thermo Life Science, Dreieich
Thermomixer Comfort	Eppendorf, Hamburg
Thermofluor Gerät	Stratagene, Waldbronn
Tiefkühlschrank Labcool	Sanyo, Bad Nenndorf
Ultra-low Tiefkühlschrank	New Brunswick, Nürtingen
Ultraschallbad	Bandelin, Berlin
Vortexer	VWR, Darmstadt
Wasserbad	Labora, München
Zentrifuge 5424	Eppendorf, Hamburg
Zentrifuge 5424R	Eppendorf, Hamburg
Zentrifuge 5810R	Eppendorf, Hamburg

4.2.2. Methoden

4.2.2.1. Mikrobiologische Methoden

4.2.2.1.1. Bakterienkulturen

Lagerung:

Die BL21(DE3)*pLysS*, BL21(DE3) und NEB5 α -Zellen wurden bei -80 °C gelagert. Diese wurden auf LB-Agar Platten nach der Transformation bis zu wenigen Wochen bei 4 °C im Dunkeln gelagert. Die Zellpellets aus Expressionskulturen wurden bei -20 °C gelagert.

Übernachtskultur für die Miniplasmidpräparation:

Für die Miniplasmidpräparation wurden Suspensionsschüttelkulturen des NEB5 α -Stamms mit je 5 mL LB-Medium in einem Kulturröhrchen bei 37 °C und 200 rpm unter aeroben Bedingungen über Nacht angesetzt. Die Selektion konnte durch Antibiotika ($C_{\text{END}} = 100 \mu\text{g/mL}$) Zugabe erreicht werden. Zum Animpfen des LB-Mediums wurde eine Bakterienkultur von der Agar-Platte entnommen und in das Medium überführt. Dabei wurde unter sterilen Bedingungen gearbeitet

4.2.2.1.2. Transformation chemisch kompetenter *E. Coli* Zellen

Die chemisch kompetenten Zellen wurden auf Eis aufgetaut und mit etwa 100 ng Plasmid-DNA versetzt. Anschließend erfolgte eine Inkubation auf Eis für 20-30 min. Danach wurde ein Hitzeschock bei 37 °C für 2 min durchgeführt, woraufhin erneut eine Inkubation auf Eis für 5 min erfolgte. Als Nächstes wurde 950 μL LB- Medium zu den Zellen pipettiert und bei 37 °C unter Schütteln für 1 h inkubiert. Hiernach wurde für 5 min bei 15000 rpm zentrifugiert und 900 μL des Überstands verworfen. Das Pellet wurde im Restmedium resuspendiert und auf einer Agar-Platte mit dem entsprechenden Antibiotikum ausgestrichen. Nach Inkubation über Nacht bei 37 °C, erfolgte die Lagerung bei 4 °C. Für die XPA-Expression wurde der *E.coli* Expressionsstamm BL21(DE3)*pLysS* verwendet, wodurch neben Carbenicillin zusätzlich das Antibiotikum Chloramphenicol benutzt wurde. Für die RPA-Expression wurde der *E.coli* Expressionsstamm BL21(DE3) verwendet.

4.2.2.1.3. Heterologe Proteinexpression in *E.coli*

XPA:

Nach erfolgreicher Transformation des XPAfl-pPSG-IBA35 Vektors in den *E. coli* BL21(DE3)*pLysS* Expressionsstamm wurde eine 500 mL LB-Übernachtskultur mit dem entsprechenden Antibiotikum angesetzt, um das selektive Wachstum resistenter Bakterienstämme zu gewähren. Zur Inokulation des LB-Mediums wurden Bakterienkolonien von der entsprechenden Agar-Platte entnommen und in das LB-Medium überführt. Die Übernachtskultur wurde zuerst 1 h bei 37 °C und 200 rpm inkubiert. Danach wurde 12 h bei 37 °C und 0 rpm inkubiert. Zur Proteinexpression wurde je 1 L LB-Medium (insgesamt 10 x 1 L) mit je 45 mL Übernachtskultur und dem entsprechenden Antibiotikum (1:1000) angeimpft. Das angeimpfte Medium wurde bei 37 °C und 200 rpm bis zu einer OD₆₀₀ von 0.6 inkubiert. Anschließend wurde mit 0.5 mM IPTG und 10 µM ZnCl₂ induziert. Dabei diente ZnCl₂ der Unterstützung zum Aufbau des Zinkfinger-Motivs von XPA. Nach 4 h bei 37 °C und 200 rpm Inkubation wurden die Bakterien bei 4 °C und 8000 rpm für 8 min zentrifugiert. Nachdem der Überstand verworfen wurde, wurden die Pellets bei -20 °C gelagert.

RPA:

Nach erfolgreicher Transformation des p11d-tRPA Vektors in den *E. coli* BL21(DE3) Expressionsstamm wurden je 1 L LB-Medium (insgesamt 5 x 1 L) mit einer Bakterienkolonie und dem entsprechenden Antibiotikum zur Proteinexpression angeimpft (1:1000), um das selektive Wachstum resistenter Bakterienstämme zu gewährleisten. Das angeimpfte Medium wurde über Nacht für 12 h bei 37 °C und 0 rpm inkubiert. Anschließend wurde das Medium bei 37 °C und 200 rpm bis zu einer OD₆₀₀ von 0.6 inkubiert, woraufhin mit 0.3 mM IPTG induziert wurde. Nach 3 h bei 37 °C und 200 rpm Inkubation wurden die Bakterien bei 4 °C und 8000 rpm für 8 min zentrifugiert. Nachdem der Überstand verworfen wurde, wurden die Pellets bei -20 °C gelagert.

4.2.2.1.4. Zellaufschluss mit Ultraschall

Der Ultraschall Homogenisator (Sonoplus HD 2070, Bandelin, Berlin) diente zum Aufschluss der Bakterienzellen. Dabei wurden die Bakterienpellets (aus 2.5 L Flüssigkultur) auf Eis aufgetaut und im entsprechenden Puffer A (20 mL) der verwendeten Säule, sowie mit einer halben (RPA) bzw. viertel (XPA) Tablette Protease Inhibitor *Complete*® EDTA-frei (Roche, Mannheim) resuspendiert. Danach wurde das resuspendierte Pellet mit Lysozym (c_{END} = 0.5 mg/mL) für 30 min auf Eis inkubiert.

Anschließend erfolgte der Aufschluss auf Eis mit 10 Zyklen für je 30 s bei maximaler Energie und einer Pulseinstellung von 5. Bei der XPA Überexpression wurde der Suspension vor der Zentrifugation Benzonase (10 U/mL) zugegeben, um die Viskosität der Lösung zu reduzieren. Nach dem Zellaufschluss wurde die Suspension bei 18000 rpm und 4 °C für 45 min zentrifugiert. Das erhaltene Lysat wurde im weiteren Verlauf für die Proteinreinigung verwendet.

4.2.2.2. Molekularbiologische Methoden

4.2.2.2.1. DNA Isolierung - Mini-Plasmidpräparation

Zur Isolierung von kleinen Plasmid-DNA Mengen wurde das *GeneJET* Plasmid Miniprep Kit (Thermo Scientific, Dreieich) nach Herstellerangaben verwendet.

4.2.2.2.2. Klonierung mittels StarGate®-Technologie

Die StarGate®-Technologie ermöglichte die Klonierung des *xpafl* Gens in den pPSG-IBA35 Vektor mit einem N-terminalen His-tag. Der Donorvektor IBA10 mit dem *xpafl* Gen wurde bereits von *Sandra Koch* hergestellt und verwendet. Im nächsten Schritt wurde die Transferreaktion des Donorvektors in den Akzeptorvektor durchgeführt. Dazu wurde der Akzeptorvektor (20 ng) mit dem Donorvektor (24 ng) für 1 h bei 30 °C mit der *StarCombinase*TM inkubiert. Anschließend wurden 15 µL des Reaktionsansatzes in *E. coli* BL21(DE3)*pLysS* Zellen transformiert. Zur Isolierung des Destinationsvektors wurde eine Mini-Plasmidpräparation durchgeführt. Nach Isolierung des Destinationsvektors wurde dieser in NEB5α Zellen transformiert und erneut eine Mini-Plasmidpräparation durchgeführt.

4.2.2.2.3. Restriktionsverdau

Der Restriktionsverdau wurde zur Überprüfung des Klonierungsprodukts (Abschnitt 4.2.2.2.2.) verwendet. Die genutzten Restriktionsenzyme HindIII (Erkennungsspaltstelle A/AGCTT) und XbaI (Erkennungsspaltstelle T/CTAGA) sind Endonukleasen und unterscheiden sich in ihrer Erkennungssequenz. Der Restriktionsverdau ermöglicht somit die Entstehung verschieden langer, palindromischer Sequenzen, die sich in einem Agarosegel analysieren lassen. In Tabelle 9 ist der Restriktionsansatz aufgelistet. Der Verdau wurde bei 37 °C für 2 h durchgeführt. Die Inaktivierung

der Enzyme erfolgte bei 70 °C für 5 min. Die Restriktionsfragmente wurden anschließend *via* DNA-Agarosegelelektrophorese untersucht.

Tabelle 9: Zusammensetzung des Restriktionsansatzes.

Zusammensetzung	Menge
Plasmid-DNA (pPSG-IBA35-XPAfl)	4 µL (ca. 1000 ng)
10x NEB 2.1-Puffer	1 µL
HindIII-Restriktionsenzym	0.22 µL
XbaI-Restriktionsenzym	0.22 µL
H ₂ O	4.56 µL

4.2.2.2.4. DNA-Agarosegelelektrophorese

Die Agarosegelelektrophorese wurde zur Auftrennung von Nukleinsäure-Fragmenten, z.B. eines Restriktionsverdaus, benutzt. Diese werden nach ihrer Größe/Konformation aufgetrennt und mit einem DNA-Standard zur Größenbestimmung verglichen. Die Proben für die Agarosegelelektrophorese wurden vor dem Beladen mit je 2 µL 5x-Agarose-Auftragspuffer versetzt. Die Taschen des Gels wurden mit 10 µL Probe bzw. 7 µL DNA-Standard beladen. Dabei wurde ein 1%-iges Agarosegel verwendet und eine Spannung von 100 V für 1 h angelegt. Durch Anlegen einer Spannung wandert die negativ geladene DNA zur Anode, wobei kleinere DNA-Fragmente die Agarosematrix schneller passieren. Zum Nachweis der DNA wurde Ethidiumbromid, welches in die DNA interkaliert, benutzt. Dieses wurde bereits beim Gießen des Gels hinzugegeben. Ethidiumbromid konnte mit ultravioletterm Licht ($\lambda_{\text{abs}} = 254 \text{ nm}$, $\lambda_{\text{em}} = 595 \text{ nm}$) detektiert werden. Tabelle 10 zeigt die Zusammensetzung eines 1%-igen Agarosegels auf.

Tabelle 10: Zusammensetzung eines 1%-igen Agarosegels für ein 50 mL Gel.

Lösungen	Menge
TAE-Gelpuffer	50 mL
Agarose	0.5 g
Ethidiumbromid	7 µL (1 mg/mL in ddH ₂ O)

4.2.2.2.5. DNA-Sequenzierung

Zur Überprüfung der richtigen Sequenz des Destinationsvektors wurde die Sequenzierung durch die Firma GATC (Konstanz) durchgeführt. Die dabei erhaltenen Sequenzen wurden mit dem Programm *CLC-Sequence Viewer* ausgewertet, um die Übereinstimmung mit der theoretischen Sequenz zu prüfen.

4.2.2.3. Proteinchemische Methoden

4.2.2.3.1. XPA_{fl} Reinigung

Die Chromatographieschritte zur XPA_{fl} Reinigung sind in Tabelle 11 aufgelistet. Zur Aufreinigung wurde ein Äkta System verwendet, welches eine Flüssigkeitschromatographie zur schnellen Aufreinigung von Proteinen ist (FPLC, *fast protein liquid chromatography*).

Tabelle 11: Chromatographieschritte der XPA_{fl} Reinigung.

Chromatographieschritte	Chromatographiesäule
1.	Ni-NTA Affinitätschromatographie
2.	MonoQ Anionenaustauschchromatographie
3.	Heparin Affinitätschromatographie
4.	Gelfiltrationschromatographie (SEC)

1. Ni-NTA Affinitätschromatographie:

Bei der Ni-NTA Affinitätschromatographie können Proteine mit einem Polyhistidin-tag (6xHis-tag) an Metallchelate (Ni-NTA) gebunden werden, wodurch sich ein stabiler oktaedrischer Chelatkomplex ausbildet, da die Histidin Reste an das Metallion koordinieren. Zum Schluss kann das Protein mit Imidazol eluiert werden, da Imidazol kompetitiv an die Bindungsstellen des His₆-tags gebunden wird.^[152]

Zur Proteinaufreinigung von XPA_{fl}, welches einen N-terminalen His₆-tag besitzt, wurde eine 5 mL HisTrap™ Ni-NTA-Säule (GE Healthcare, Garching), die mit Ni-NTA Puffer A äquilibriert wurde, verwendet. Dabei wurde das Lysat über einen *Superloop* (2 x 40 mL) mit Ni-NTA Puffer A und einer Flussrate von 2 mL/min auf die Säule aufgetragen. Der Ni-NTA Puffer enthält 30 mM Imidazol, um unspezifische Interaktionen von anderen Proteinen zu vermeiden. Nach Auftrag des Lysats wurde

mit 100 % Ni-NTA Puffer A für 5 Säulenvolumen (CV) gewaschen. Anschließend konnte das Protein mit einem linearen Gradienten von 0% auf 100% Ni-NTA Puffer B in 8 CV bei einer Flussrate von 1 mL/min und einer Fraktionsgröße von 2.5 mL eluiert werden. Die Protein enthaltenden Fraktionen wurden mittels SDS-PAGE (SDS-Polyacrylamidgelelektrophorese) identifiziert.

2. MonoQ Anionenaustauschchromatographie:

Die MonoQ ist ein Anionenaustauscher und besitzt ein positiv geladenes Säulenmaterial. Da das XPA Protein einen negativen isoelektrischen Punkt (-6.57) bei einem pH-Wert von 7.5 besitzt, kann dieses an die Säule binden. Das Protein kann durch eine erhöhte Salzkonzentration eluiert werden.^[153]

Die XPA enthaltenden Fraktionen aus dem Ni-NTA Reinigungsschritt wurden im MonoQ Puffer A im Verhältnis 1:2 verdünnt und über einen *Superloop* mit einer Flussrate von 1 mL/min auf eine 1 mL MonoQTM 5/50 GL (GE Healthcare, Garching), die mit MonoQ Puffer A äquilibriert wurde, aufgetragen. Danach wurde mit MonoQ Puffer A für 5 CV gewaschen. Die Proteinelution erfolgte bei einem linearen Gradienten von 0% auf 100% MonoQ Puffer B in 15 CV bei einer Flussrate von 1 mL/min und einer Fraktionsgröße von 1 mL. Protein enthaltende Fraktionen wurden mittels SDS-PAGE analysiert.

3. Heparin Affinitätschromatographie:

Um einen höheren Reinheitsgrad von XPA zu erreichen, wurde als nächstes die 1 mL HiTrapTM Heparin HP Säule (GE Healthcare, Garching) verwendet. Die Heparin Säule wird sowohl zur Affinitätschromatographie für DNA-bindende Proteine, als auch als Kationenaustauscher verwendet. Auch hier erfolgt die Elution des Proteins durch eine Erhöhung der Salzkonzentration.^[154] XPA kann durch die starke DNA-Bindungsaffinität an die Heparin-Säule binden, obwohl es im Heparin-Puffer A eine negative Ladung aufweist.

XPA enthaltende Fraktionen aus der MonoQ Anionenaustauschchromatographie wurden mit Heparin Puffer A im Verhältnis 1:7 verdünnt und über einen *Superloop* mit einer Flussrate von 1 mL/min auf die Säule, die mit 5% Heparin Puffer B äquilibriert wurde, aufgetragen. Danach wurde mit 5% Heparin Puffer B für 5 CV gewaschen. Die Proteinelution erfolgte bei einem linearen Gradienten von 5% auf 100% Heparin Puffer B in 35 CV bei einer Flussrate von 1 mL/min und einer Fraktionsgröße von 1 mL. Protein enthaltende Fraktionen wurden mittels SDS-PAGE analysiert.

4. Gelfiltrationschromatographie:

Als letztes wurde die 120 mL Gelfiltrationssäule HiLoadTM 16/600 Superdex 75 *prep grade* (GE Healthcare, Garching) verwendet, um weitere Verunreinigungen und Aggregate abzutrennen. Dabei diffundieren kleinere Moleküle in die poröse Matrix der Säule, wodurch die Retentionszeit im

Vergleich zu größeren Molekülen vergrößert wird. Somit wandern die größeren Moleküle schneller durch die Säule und eluieren als erstes.^[155]

Aus der Heparin Aufreinigung erhaltene XPA Fraktionen wurden zunächst mit einem Amicon Röhrchen (15 mL, 10 kDa, Millipore) durch Zentrifugation mit dem Kristallisationspuffer II umgepuffert und auf ca. 2 mL einkonzentriert. Danach wurde die Proteinprobe durch einen 5 mL *Sampleloop* mit einer Flussrate von 1 mL/min auf die Gelfiltrationssäule, die bereits mit Kristallisationspuffer II äquilibriert wurde, aufgetragen. Dabei betrug die Fraktionsgröße 2 mL. Zum Schluss wurden die sauberen Protein enthaltenden Fraktionen mittels SDS-PAGE analysiert, vereint und mit einem Amicon Röhrchen (15 mL, 10 kDa, Millipore) durch Zentrifugation auf ca. 8-10 mg/mL einkonzentriert. Anschließend wurde das Protein in flüssigem Stickstoff schockgefroren und bei -80 °C gelagert.

4.2.2.3.2. RPA Reinigung

Die Chromatographieschritte zur RPA Reinigung sind in Tabelle 12 aufgelistet. Dabei wurde nach einem Protokoll von *Wold et al.* verfahren.^[156] Zur Aufreinigung wurde ein Äkta System verwendet.

Tabelle 12: Chromatographieschritte der RPA Reinigung.

Chromatographieschritte	Chromatographiesäule
1.	Cibacron TM Blue Affinitätschromatographie
2.	Hydroxylapatit (HAP) Affinitätschromatographie
3.	MonoQ Anionenaustauschchromatographie

1. CibacronTM Blue Affinitätschromatographie:

Bei der CibacronTM Blue Affinitätschromatographie ist der Farbstoff CibacronTM Blue F3G-A als Ligand an der Säulenmatrix gebunden. Dadurch fungiert die Säule als starker Kationentauscher und DNA-interagierende Proteine können ebenfalls an die Säule binden.^[157-158]

Zur Proteinaufreinigung von RPA, welches eine starke DNA Bindungsaffinität besitzt, wurde eine 8 mL (1x 5 mL und 3x 1mL) HiTrapTM Blue HP Säule (GE Healthcare, Garching), die mit 5% HI Puffer B äquilibriert wurde, verwendet. Dabei wurde das Lysat über einen *Superloop* (40 mL) mit 5% HI-Puffer B (95% HI-Puffer A) und einer Flussrate von 1 mL/min auf die Säule aufgetragen. Nach dem Auftrag des Lysats wurde mit 5% HI-Puffer B für 5 CV gewaschen. Als nächstes wurde mit 80% HI-Puffer B und danach mit 100% HI-Puffer C gewaschen. Anschließend konnte das Protein mit 100%

HI-Puffer D (1.5 mM NaSCN) eluiert werden. Die Waschfraktionen, als auch die Elutionsfraktionen, wurden mit einer Fraktionsgröße von 8 mL gesammelt. Protein enthaltende Fraktionen wurden mittels SDS-PAGE identifiziert.

2. HAP Affinitätschromatographie:

Bei der HAP Säule handelt es sich sowohl um einen Kationenaustauscher als auch um eine Calcium-Affinitätschromatographie. Die Aminogruppen des Proteins können mit den negativen Phosphatgruppen des HAP-Trägermaterials ($(\text{Ca}_5(\text{PO}_4)_3\text{OH})_2$) interagieren. Weiterhin können Proteine durch ihre Carboxylgruppen mit Calcium interagieren. Die Elution kann mit Phosphat im Puffer erfolgen.^[159]

Aus der CibacronTM Blue Affinitätschromatographie erhaltene RPA Fraktionen, wurden mit dem HI-Puffer E im Verhältnis 1:2 verdünnt und über einen *Superloop* mit einer Flussrate von 1 mL/min auf eine selbstgepackte 3.5 mL HAP Säule, die mit HI-Puffer E äquilibriert wurde, aufgetragen. Danach wurde mit HI-Puffer E für 5 CV gewaschen. Die Protein-Elution erfolgte bei 100% HI-Puffer F, einer Flussrate von 1 mL/min und einer Fraktionsgröße von 1 mL. Zum Schluss wurde die Säule mit HI-Puffer G sowie Lagerungspuffer 1 gewaschen und im Lagerungspuffer 2 gelagert. Protein enthaltende Fraktionen wurden mittels SDS-PAGE analysiert.

3. MonoQ Anionenaustauschchromatographie:

Zur weiteren Aufreinigung von RPA wurde ein Stufengradient für die MonoQ Anionenaustauschchromatographie benutzt. RPA enthaltende Fraktionen (aus dem HAP Affinitätschromatographie-Schritt) wurden mit HI-Puffer A im Verhältnis 1:2 verdünnt und über einen *Superloop* mit einer Flussrate von 1 mL/min auf die Säule, die mit 5% HI-Puffer B äquilibriert wurde, aufgetragen. Danach wurde mit 5%, 10% und mit 20% HI-Puffer B für jeweils 5 CV gewaschen (Stufengradient). Die Elution des Proteins erfolgte mit einem linearen Gradienten von 20% auf 40% HI-Puffer B in 10 CV bei einer Flussrate von 1 mL/min und einer Fraktionsgröße von 1 mL. Danach wurde die Säule mit 40% und 100% HI-Puffer B gewaschen. Protein enthaltende Fraktionen wurden mittels SDS-PAGE analysiert. Zum Schluss wurden die reinsten Protein enthaltenden Fraktionen vereint und mit einem Amicon Röhrchen (15 mL, 30 kDa, Millipore) durch Zentrifugation mit dem Kristallisationspuffer II umgepuffert und auf ca. 8-12 mg/mL einkonzentriert. Anschließend wurde das Protein in flüssigem Stickstoff schockgefroren und bei -80 °C gelagert.

4.2.2.3.3. Bestimmung der Proteinkonzentration

Die Konzentration von Proteinen wurde mit dem NanoDrop Spektrophotometer ND-1000 (Peglab, Erlangen) bestimmt. Dabei wurde die Konzentration des Proteins durch die gemessene Absorption bei 280 nm bestimmt, da die aromatischen Aminosäuren Phenylalanin, Tryptophan und Tyrosin ihr Absorptionsmaximum bei 280 nm haben. Bei einem Extinktionswert von 1 entspricht dies einer Konzentration von ca. 1 mg/mL. Für genauere Messwerte kann der jeweilige Extinktionskoeffizient des Proteins mit eingerechnet werden. Weiterhin kann die Reinheit der Proteinprobe durch diese photometrische Messung bestimmt werden, um DNA-Kontaminationen auszuschließen. Ein Wert unter 1 des Verhältnisses von 260 nm : 280 nm zeigt eine reine Proteinprobe auf.^[144]

4.2.2.3.4. Konzentrierung der Proteinlösung

Für die Aufreinigungsschritte oder für weitere biochemische Studien war das Aufkonzentrieren der Proteinlösung notwendig. Dabei wurden die Protein enthaltenden Fraktionen mit einem Amicon Ultra-15 und Amicon Ultra-4 (Merck Millipore, Darmstadt) durch Zentrifugation bei 3220 xg und 4 °C aufkonzentriert. Für kleinere Volumina wurde ein Amicon Ultra-0.5 mL verwendet (14674 xg, 4 °C). Für XPA wurde ein Zentrifugenfilter mit einer Ausschlussgröße von 10 kDa und für RPA von 30 kDa verwendet.

4.2.2.3.5. SDS-PAGE

Zur Überprüfung der Reinigung und zur Auftrennung der Proteine nach ihrem Molekulargewicht wurde die SDS-PAGE unter denaturierenden Bedingungen durchgeführt. Durch Anlegen eines elektrischen Feldes können die Proteine zur Anode wandern und lassen sich dadurch elektrophoretisch auftrennen. Zudem kann SDS die Eigenladung von Proteinen überdecken, wodurch die Proteine negativ geladen sind und nur nach ihrer Größe elektrophoretisch aufgetrennt werden. Dabei wandern kleinere Proteine schneller durch die Poren des Polyacrylamidgels. Durch einen Proteinstandard kann das Molekulargewicht der einzelnen Proteine bestimmt werden.^[160]

Die Proteinproben wurden vor dem Beladen mit einer geeigneten Menge an 5x-SDS-Auftragspuffer versetzt und bei 95 °C für 5 min denaturiert. Die Taschen des Gels wurden mit 5-20 µL Probe bzw. 5 µL Protein-Standard beladen. Dabei wurde ein 15% SDS-Gel verwendet und die Proteine in einem elektrischen Feld für 30 min bei 300 V oder 60 min bei 180 V aufgetrennt. Danach wurde das Gel mit

einer Coomassie-Lösung II (600 µL) und 10 mL Färbelösung II für 15 min gefärbt und mit Wasser für 30 min entfärbt, um die Banden der Proteine und des Markers sichtbar zu machen.

Tabelle 13 zeigt die Zusammensetzung eines 15% SDS-Gels auf. Die Lagerung der gegossenen Gele erfolgte bei 4 °C.

Tabelle 13: Zusammensetzung für fünf 15%-ige SDS-Gele.

Reagenzien	15% Trenngel	5% Sammelgel
ddH ₂ O	4.6 mL	6.8 mL
Rotiphoese Gel 30	10 mL	1.7 mL
Tris-HCl 1.5 M, pH 8.8	5 mL	-
Tris-HCl 1 M, pH 6.8	-	1.25 mL
10% SDS-Lösung	200 µL	100 µL
10% APS-Lösung	200 µL	100 µL
TEMED	20 µL	10 µL

4.2.2.5. Biochemische Methoden

4.2.2.5.1. EMSA

Mit EMSA Studien können DNA-Protein-Interaktionen untersucht werden. Durch die DNA-Protein-Interaktion wird das Laufverhalten der DNA im Vergleich zu freier DNA in einem elektrischen Feld verändert, wodurch die Interaktion nach Anfärbung der DNA nachgewiesen werden kann.^[161]

Kleines 4.5% natives Polyacrylamidgel:

Die dsDNA (10 pmol pro Spur) wurde mit steigender Proteinkonzentration bei verschiedenen Zeiten und Temperaturen (abhängig vom Experiment) inkubiert. Für das Bindungsexperiment wurde ein 10 µL Reaktionsansatz in Kristallisationspuffer II gewählt, welcher mit 2 µL 5x EMSA-Auftragspuffer versetzt wurde. Anschließend wurden die Taschen des 4.5% nativen Gels mit 10 µL Probe beladen. Danach wurde ein elektrisches Feld für 45 min bei 100 V und 4 °C angelegt (Laufpuffer 0.25x TBE). Danach wurde das Gel mit einer 1x TBE-Lösung (50 mL) und 5 µL SYBR Green I (1:10000, Sigma Aldrich, Deisenhofen) für 20 min im Dunkeln und bei RT gefärbt, um die Banden der DNA sichtbar zu machen. Die Detektion erfolgte bei einer Absorption mit 460 nm und einer Emission mit 515 nm über das *Image Reader* Gerät. Tabelle 14 beschreibt das Pipettierschema der Gele.

Tabelle 14: Zusammensetzung für fünf 4.5%-ige native TBE-Gele.

Reagenzien	Volumen
ddH ₂ O	23.25 mL
Rotiphorese 30	4.5 mL
10x TBE	1.5 mL
Glycerin 100%	0.75 mL
10% APS-Lösung	225 µL
TEMED	15 µL

Großes 3.5% natives Polyacrylamidgel:

Die dsDNA (10 pmol/µL, 2-4 µL) wurde mit steigender Proteinkonzentration bei verschiedenen Zeiten und Temperaturen (abhängig vom Experiment) inkubiert. Für das Bindungsexperiment wurde ein 20 µL Reaktionsansatz in Kristallisationspuffer II gewählt, welcher mit 4 µL 5x EMSA-Auftragspuffer versetzt wurde. Anschließend wurden die Taschen des 3.5% nativen Gels mit 20 µL Probe beladen. Danach wurde ein elektrisches Feld für ca. 3 h bei 30 mA (konstant), 1000 V, 300 W und 7 °C angelegt (Laufpuffer 1x TBE). Danach wurde das Gel mit einer 1x TBE-Lösung (200 mL) und 20 µL SYBR Green I (1:10000, Sigma Aldrich, Deisenhofen) für 20 min im Dunkeln und bei RT gefärbt, um die DNA-Banden sichtbar zu machen. Die Detektion erfolgte bei einer Absorption mit 460 nm und einer Emission mit 515 nm am Gerät *Image Reader*. Die Herstellung der Gele erfolgte wie in Tabelle 15 beschrieben.

Tabelle 15: Zusammensetzung eines großen 3.5% nativen TBE-Gels.

Reagenzien	Volumen
ddH ₂ O	40.42 mL
Rotiphorese 30	5.83 mL
10x TBE	2.5 mL
Glycerin 100%	1.25 mL
10% APS-Lösung	375 µL
TEMED	25 µL

4.2.2.5.2. Gelfiltrationsassay

Der Gelfiltrationsassay dient zur Isolierung und zum Nachweis eines XPA/RPA/DNA-Komplexes. Dafür wurden verschiedene Verhältnisse des Komplexes getestet (4:4:1, 1:2:2, 1:1:1 für XPA:RPA:DNA). Die verwendete Gelfiltrationssäule beruht auf dem SEC-Prinzip, wodurch der Komplex im Gegensatz zu den einzelnen Komponenten als erstes eluiert werden sollte, da dieser im Vergleich zu den einzelnen Proteinen und der DNA größer ist.

Die Proteine und DNA, welche im Kristallisationspuffer II vorlagen, wurden in einem Reaktionsansatz von insgesamt 200 µL zusammen pipettiert und für 5 min bei 15000 rpm und bei 4 °C zentrifugiert. Dabei wurde nach Optimierung des Assays der Kristallisationspuffer II anstatt I verwendet. Nach der Zentrifugation wurde der Überstand durch einen 500 µL *Sampleloop* mit einer Flussrate von 0.5 mL/min bei 4 °C auf eine 24 mL Gelfiltrationssäule (Superdex™ 200 Increase 10/300 GL, GE Healthcare), die bereits mit Kristallisationspuffer II äquilibriert wurde, aufgetragen. Dabei wurden 0.5 mL Fraktionen für 60 min gesammelt. Durch die Absorption bei 260 nm und 280 nm wurden Protein und DNA enthaltende Fraktionen identifiziert.

Anschließend wurden diese mittels SDS-PAGE analysiert, um den XPA/RPA/DNA-Komplex enthaltende Fraktionen zu identifizieren. Dafür wurden 20 µL der gesammelten Fraktion mit je 4 µL 5x SDS-Auftragspuffer versetzt und bei 95 °C für 5 min denaturiert. Die Taschen des 15% SDS-Gels wurden mit 20 µL Probe bzw. 5 µL Protein-Standard beladen. Dabei wurden die Proteine in einem elektrischen Feld für 60 min bei 180 V aufgetrennt. Danach wurde das Gel mit der Coomassie-Lösung II (600 µL) und Färbelösung II (10 mL) für 15 min gefärbt und mit Wasser für 30 min entfärbt, um die Banden der Proteine und des Markers sichtbar zu machen.

Weiterhin wurden gesammelte Fraktionen mittels nativer Gelelektrophorese analysiert, um DNA enthaltende Fraktionen zu identifizieren und um Protein-DNA Interaktionen nachweisen zu können. Dafür wurden 10 µL der gesammelten Fraktion mit je 2 µL 5x EMSA-Auftragspuffer versetzt und die Taschen des 4.5% nativen Gels mit 10 µL Probe beladen. Anschließend wurde ein elektrisches Feld für 45 min bei 100 V und 4 °C angelegt. Danach wurde das Gel mit einer 1x TBE-Lösung (50 mL) und 5 µL SYBR Green I für 20 min im Dunkeln und bei RT gefärbt, um die Banden der DNA sichtbar zu machen. Die Detektion erfolgte bei einer Absorption mit 460 nm und einer Emission mit 515 nm am Gerät *Image Reader*. Daneben wurden XPA und RPA einzeln, zusammen und mit DNA, sowie die DNA ohne Protein mit dem Gelfiltrationsassay, wie oben beschrieben getestet, um die einzelnen Retentionszeiten/-volumina im Vergleich zum Komplex bestimmen zu können. Außerdem wurde die Stabilität des Komplexes getestet, indem die Komplex enthaltende Fraktion einkonzentriert und bei 4 °C, sowie bei -80 °C gelagert wurde, um dann erneut auf die Gelfiltrationssäule aufgetragen

wurden. Ferner wurde ein Gelfiltrationsstandard von Biorad, München (Art-Nr. 151-1901), welcher die Proteine Thyroglobulin, γ -Globulin, Ovalbumin, Myoglobin und Vitamin B12 enthielt, nach Herstellerangaben verwendet. Die einzelnen Retentionszeiten/-volumina der jeweiligen Proteine geben Rückschlüsse auf das molekulare Gewicht. Mit diesem Standard sollte das molekulare Gewicht des XPA/RPA/DNA-Komplexes bestimmt werden. Hierfür wurde der Standard in 500 μ L Kristallisationspuffer II gelöst und 100 μ L davon wurden mit einem 500 μ L *Sampleloop* und einer Flussrate von 0.5 mL/min auf die Gelfiltrationssäule aufgetragen. Die Protein enthaltenden Fraktionen wurden durch die Absorption bei 280 nm identifiziert und mittels SDS-PAGE analysiert.

4.2.2.5.3. Thermofluor

Der Thermofluor Assay ermöglicht optimale Pufferbedingungen für ein Protein zu finden. Dadurch können unterschiedliche Salze, pH-Werte und Additive getestet werden, um die Stabilität eines Proteins zu testen. Die Methode beruht auf der Exposition von hydrophoben Resten des Proteins im Verlauf der Denaturierung (durch Erhöhung der Temperatur), die durch den Farbstoff SYPRO Orange detektiert werden kann. Denn der Farbstoff kann an die freigesetzten hydrophoben Bereiche des Protein. Dabei kann eine Zunahme der Fluoreszenz detektiert werden, da die Fluoreszenz des Farbstoffs bei Interaktionen mit hydrophoben Bereichen steigt. Somit wird die Emissionskurve des Farbstoffs in Abhängigkeit der Temperatur gemessen. Umso höher die Denaturierungstemperatur des Proteins ist, desto höher ist die Proteinstabilität in diesem Puffer.^[162]

Für den Thermofluor Assay wurden 96-*well* Platten und eine *Real-Time* PCR Maschine (Biorad CFX Touch, Biorad, München) von der Hopfner Gruppe (LMU München) verwendet. Zudem wurde nach dem Protokoll von *Boivin et al.* verfahren.^[162] Dafür wurden 2 μ L Proteinlösung ($c_{\text{END}} = 1.6 \mu\text{M}$) und 2 μ L SYPRO Orange ($c_{\text{END}} = 5\times$, Sigma Aldrich, Deisenhofen) zum zu testenden Puffer gegeben (Reaktionsansatz pro *well* = 25 μ L) und die Temperatur von 5 °C auf 95 °C mit einem Gradienten von 1 °C/min erhöht. Die Extinktion für den Farbstoff betrug 485 nm und die Emission 530 nm. Die erhaltenen Datensätze wurden mit dem Microsoft Office Programm Excel ausgewertet. Für die Testbedingungen wurde sowohl eine Platte nach *Boivin et al.* (Platte: *Buffer Screen*), als auch weitere Bedingungen getestet.^[162] Im Anhang befinden sich die unterschiedlichen Bedingungen.

4.2.2.6. Proteinkristallisation

Die Kristallisationsstudien des XPA/RPA/DNA-Komplexes wurden teilweise in enger Zusammenarbeit mit Prof. Caroline Kisker und Dr. Jochen Kuper (Würzburg) durchgeführt. Proteinkristalle können die Aufklärung der dreidimensionalen Proteinstruktur *via* Röntgenstrukturanalyse ermöglichen. Proteinkristalle entstehen beim Übergang einer gesättigten Lösung in eine feste Phase, welches durch Präzipitate wie Salz oder organische Lösungen ermöglicht wird. Für die Studien wurde die Dampfdiffusions-Methode verwendet, wobei die Probenlösung eine geringere Konzentration als die Reagenzlösung besitzt. Durch Dampfdiffusion verringert sich das Volumen der Probelösung (des Tropfens), wobei die Konzentration steigt und es somit zur Kristallisation kommt.^[163]

Zuerst wurden initiale Bedingungen angesetzt (Index-Screen, Kisker Gruppe), um beispielsweise eine geeignete Salz-Konzentration oder pH-Wert für die Protein-Kristallisation zu finden. Dabei wurden die Proteine XPA (9.0 nmol in 35 µL Kristallisationspuffer II) und RPA (3.7 nmol in 35 µL Kristallisationspuffer) einzeln, sowie zusammen (je 2.5 nmol in 35 µL Kristallisationspuffer II) getestet. Um Proteinpräzipitate zu entfernen, wurden die Proteine zuvor 5 min bei 21130 xg zentrifugiert. Dabei wurde der Überstand für die Kristallisation verwendet. Mit dem Kristallisationsroboter (Honeybee von Genomic Solutions, USA) wurden 96 verschiedene Bedingungen getestet. Dafür wurde mit der *sitting-drop* Methode eine Tropfengröße von 150 nL bei einer Reservoirmenge von 60 µL angesetzt. Zum Schluss wurden die Kristallisationsplatten bei 20 °C gelagert.

Als Nächstes wurde versucht den XPA/RPA/DNA-Komplex, welcher im Kristallisationspuffer II vorlag, zu kristallisieren. Dafür wurde nach dem Gelfiltrationsassay der Komplex auf ca. 1 mg/mL aufkonzentriert. Auch hier wurde eine Tropfengröße von 300 nL, eine Reservoirmenge von 60 µL und die *sitting-drop* Methode verwendet. Dabei wurden vier verschiedene Screens mit je 96 Bedingungen unter Zuhilfenahme des Kristallisationsroboters angesetzt (JCSG+, Protein complex, Nucleix Suite und Index von der Kisker Gruppe). Zum Schluss wurden die Kristallisationsplatten bei 20 °C gelagert.

In späteren Kristallisationsstudien wurden verschiedene Screens von der Firma Qiagen (Classic, MPD, PACT und Nucleix) und Molecular Dimension (JCSG+) mit einem Kristallisationsroboter am Max-Planck-Institut (Prof. Elena Conti, Martinsried) angesetzt. Dabei wurde eine Tropfengröße von 100 nL, eine Reservoirmenge von 70 µL und die Dampfdiffusions-Methode bei 4 °C verwendet. Für die Kristallisationsansätze wurden sowohl der Komplex als auch die einzelnen Proteine bei verschiedenen Konzentrationen verwendet.

4.2.2.7. Negativfärbung und Kryo-Elektronenmikroskopie

Kryo-EM ermöglicht die 3D-Strukturbestimmung von Proteinen und Komplexen. Durch den verbesserten Detektor (K2 Summit von Gatan) lassen sich inzwischen Strukturen von kleineren Proteinen/Komplexen unter 200 kDa von bis zu 2.9 Å erhalten. Vorweg sollte durch Negativfärbung zum einen die Homogenität und die Partikelgröße bestimmt werden und zum anderen die optimale Konzentration getestet werden.^[164-165]

Die Negativfärbung Experimente wurden in Zusammenarbeit mit der Beckmann Gruppe (LMU, München) durchgeführt. Dafür wurde nach dem Gelfiltrationsassay der XPA/RPA/DNA-Komplex auf ca. 0.48 mg/mL einkonzentriert. Zur Bestimmung der optimalen Konzentration für die Negativfärbung wurde die Probe unverdünnt, im Verhältnis 1:10 und 1:20 verwendet. Dabei wurden 3.5 µL der Probe auf ein beglimmtes Grid, ein mit Kohlenstoff befilmtes Trägernetzchen, für 45 s aufgetragen. Danach wurde das Grid mit H₂O gewaschen und anschließend wurden 3.5 µL Uranylacetat als Kontrastmittel auf das Grid für 15 s aufgetragen. Das restliche Uranylacetat wurde dann mit Blotting Papier entfernt. Als nächstes wurden die Grids am Morgagni-Transmissions-elektronenmikroskopie (TEM 100 kV, Beckmann Gruppe) analysiert. Für die Grids, die homogene Partikel und eine homogene Färbung enthielten, wurden am Spirit (TEM 120 kV, Beckmann Gruppe) Daten gesammelt. Dieses wurde von *Charlotte Ungewickell* (Beckmann Gruppe) durchgeführt. Für die anschließende Evaluation der Datensammlungen für die 2D-Klassifikation wurden 10 Klassen durch manuelles Partikel-Picking mit dem Programm EMAN2 erstellt, welche für das automatische Partikel-Picking in Relion-1.3 verwendet wurden.

Zudem wurden Kryo-EM Messungen für den XPA/RPA/DNA-Komplex in Kooperation mit der Dietz Gruppe (TU München) durchgeführt. Dafür wurde nach dem Gelfiltrationsassay der XPA/RPA/DNA-Komplex auf ca. 1 mg/mL einkonzentriert, aliquotiert und teilweise in flüssigem Stickstoff schockgefroren. Die Proben wurden dann von *Fabian Köhler* (Dietz Gruppe) für Kryo-EM Studien verwendet. Dieser verwendete sowohl den frisch isolierten als auch den bei -80 °C gefrorenen Komplex.

4.2.2.8. Synthese von Oligonukleotiden mit einem C8-AAF-dG Schaden

4.2.2.8.1. Synthese von C8-AAF-dG

Die Synthese des C8-AAF-dG Schadens erfolgte nach *Simon et al.* (s. Abschnitt 3.2. und 3.3.)^[142]

4.2.2.8.2. Oligonukleotidsynthese

Die DNA-Synthese zum Einbau der C8-AAF-dG Modifikation wurden an einem ABI 394 DNA/RNA Synthesizer (Applied Biosystems) durchgeführt. Dafür wurden Standardprotokolle des Geräteherstellers verwendet. Dabei wurden 1 μ mol oder 200 nmol Maßstäbe zur Herstellung verwendet. Für die DNA-Synthese wurden *ultramild*-Konditionen nach Gillet *et al.*^[44] verwendet und die Kupplungszeit der Modifikation auf 30 min (2 x 15 min) verlängert. Die Standard Phosphoramidite (Ac-dC, dT) und die *ultramild*-Phosphoramidite (Pac-dA, *i*Pr-Pac-dG), sowie die *ultramild* Q-CPG-Träger wurden von den Firmen Glen Research oder Link Technologies bezogen. Als Aktivator wurde der Aktivator 42 (Sigma-Aldrich) und Acetonitril von der Firma Roth (Wassergehalt < 10 ppm oder <30 ppm) verwendet. Die Konzentration der gelösten Phosphoramidite in Acetonitril betrug 0.1 M (<10 ppm H₂O). Die Capping Lösung A und B, die Oxidationslösung und die Detritylierungslösung wurden selbst hergestellt.

4.2.2.8.3. DNA-Entschützung

Durch die DMT-*off* Synthese wurde am Ende der DNA-Synthese die terminale 5'-DMTr Gruppe abgespalten. Zur Entschützung der Oligonukleotide und zur Abspaltung von der Festphase wurden *ultramild*-Bedingungen (Lösung aus 10% Diisopropylamin und 0.25 M β -Mercaptoethanol in Methanol, 55 °C, 18 h, 650 rpm) verwendet. Nach der Entschützung wurden die Lösungen an einer SpeedVac eingeeengt.

4.2.2.8.4. Aufreinigung der Oligonukleotide durch HPLC

Zur Vorbereitung wurden die Oligonukleotide wieder in ddH₂O aufgenommen und durch einen 0.2 μ m Spritzenfilter filtriert. Für die analytische RP-HPLC wurden die Geräte 2695 *Separations Module*, 2996 *Photodiode Array Detector* (Waters, Eschborn) sowie die RP-Säule EC 250/4 Nucleodur 100-3 C18 (Macherey-Nagel, Düren) verwendet. Weiterhin wurden für die präparative RP-HPLC die Geräte 2487 *Dual λ Absorbance Detector*, 1525 *Binary HPLC Pump* (Waters, Eschborn) und die RP-Säule VP 250/10 Nucleodur 100-5 C18 (Macherey-Nagel, Düren) verwendet. Zudem konnte die aromatische Aminofluoren Einheit bei einer zweiten Wellenlänge bei 305 nm detektiert werden. Für die analytische HPLC der Oligonukleotide wurde ein Gradient von 0-30% Puffer B in 45 min und einer Flussrate von 0.5 mL/min verwendet. Für die präparative HPLC und zur Aufreinigung der kleineren

Oligonukleotide (bis zum 18mer) wurde ein Gradient von 0-30% Puffer B in 45 min und für die größeren Oligonukleotide (25mer bis 45mer) wurde ein Gradient von 0-25% in 65 min mit einer Flussrate von 5 mL/min verwendet. Die aufgereinigten Oligonukleotide wurde mittels Sep-Pak Kartuschen (Waters, Eschborn) nach dem Standardprotokoll des Herstellers entsalzt.

4.2.2.8.5. MALDI-TOF-Analyse der Oligonukleotide

Die entsalzten Oligonukleotide wurden auf einer HPA-Matrix (3-Hydroxypicolin-Säure (50 mg), 15-Krone-5 (10 µL), Ammonium-Hydrogencitrat (10 mg in 500 µL H₂O und 500 µL CH₃CN) ko-kristallisiert. MALDI-TOF Spektren wurden zur Identifizierung der richtigen Masse der Oligonukleotide mit einem Bruker-Autoflex II-Massenspektrometer und einem AnchorChip™-Target (Bruker MTP var/384) generiert.

4.2.2.8.6. Bestimmung der DNA-Konzentration

Die Konzentration der Oligonukleotide wurde mit dem NanoDrop Spektrophotometer ND-1000 (Peglab, Erlangen) bestimmt. Dabei wurde die Konzentration der gelösten Oligonukleotide durch die gemessene Absorption bei 260 nm und durch den Extinktionskoeffizienten bestimmt. Der Extinktionskoeffizient des Oligonukleotids wurde mit Hilfe des Lambert-Beerschen Gesetzes und aus der Summe der einzelnen Extinktionskoeffizienten berechnet (dA: 15.2 L mmol⁻¹ cm⁻¹, dC 7.1 L mmol⁻¹ cm⁻¹, dG 12.0 L mmol⁻¹ cm⁻¹, dT 8.4 L mmol⁻¹ cm⁻¹ und für C8-AAF-dG wurde der Extinktionskoeffizient von dG angenommen). Ein Wert zwischen 1.8 und 2.0 bei einem Verhältnis von 260 nm/280 nm spiegelt eine reine DNA-Probe ohne Protein-Kontaminationen wider.^[144]

4.2.2.8.7. Hybridisierung von DNA-Proben

Äquivalente Mengen von zwei komplementären DNA-Strängen wurden vereint und an einer SpeedVac eingengt. Anschließend wurden diese in der gewünschten Konzentration im Kristallisationspuffer I oder II angelöst, sowie zur Hybridisierung in einem Thermocycler auf 95 °C erhitzt und mit einem Gradienten von 1 °C/min auf 4 °C abgekühlt.

4.3. Ergebnisse und Diskussion

4.3.1. Synthese des C8-AAF- Phosphoramidits und DNA Festphasensynthese

Um die die Funktion eines XPA/RPA/DNA-Komplexes zu untersuchen, wurden verschiedene Oligonukleotide mit einem C8-AAF-dG Schaden synthetisiert. Für die Synthese des C8-AAF-dG Phosphoramidits (Abbildung 19) konnte eine neue, kürzere Syntheseroute etabliert werden. Diese wurde bereits im Abschnitt 3.2. und 3.3. beschrieben und veröffentlicht.

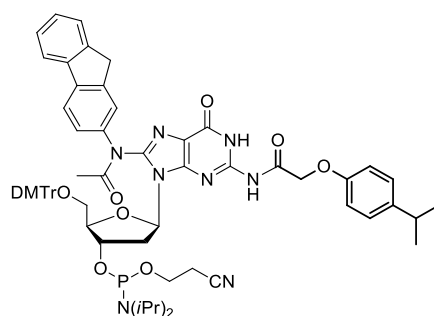
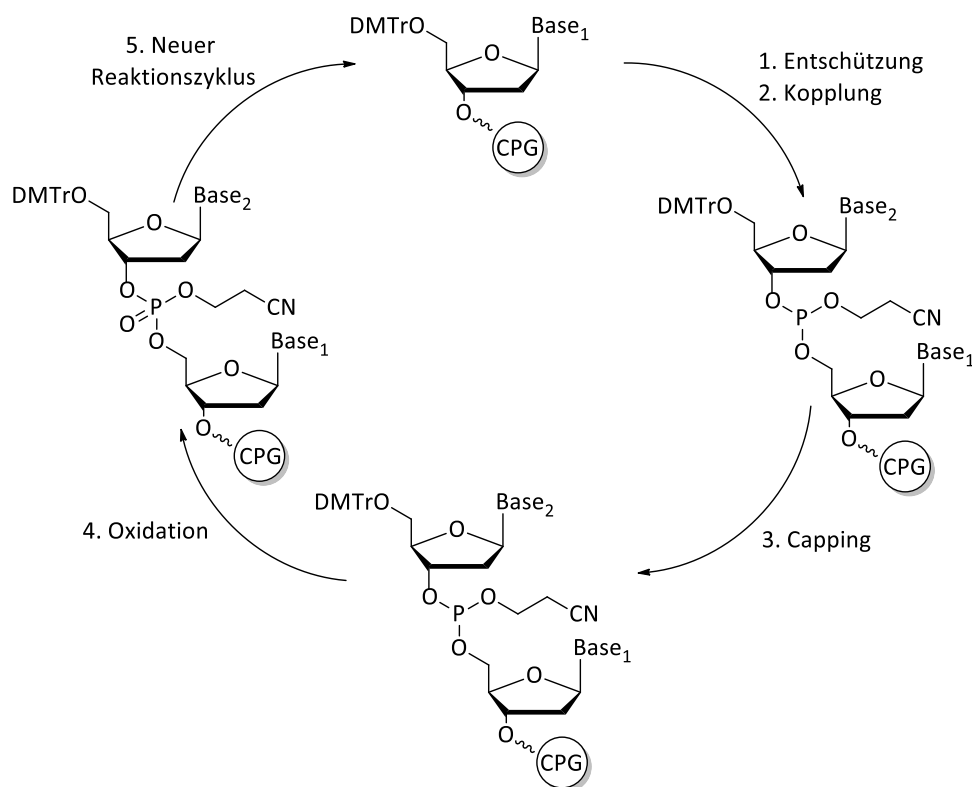


Abbildung 19: Darstellung des C8-AAF-dG Phosphoramidits.

Das synthetisierte C8-AAF-dG wurde gezielt in verschiedene Sequenzen durch automatische DNA-Festphasensynthese nach *Caruthers* erfolgreich eingebaut, welche in Schema 4 veranschaulicht ist.^[166] Dafür wurden *ultramild*-Konditionen nach *Gillet et al.* angewendet, um die Abspaltung der basenlabilen Acetylgruppe des Acetylaminofluorens während der Entschützung zu vermeiden.^[44] In der Tabelle 4 im Abschnitt 4.2.1.2. sind die hergestellten, modifizierten Oligonukleotide aufgelistet. Die Sequenzen der Oligonukleotide wurden so konzipiert, dass sich keine Sekundärstrukturen wie beispielsweise Haarnadelstrukturen ausbilden können.

Die Synthese des Oligonukleotids erfolgt von 3' nach 5' an einem festen Träger (*controlled pore glass*, CPG). Zu Beginn der Synthese ist das erste Nukleosid bereits an die CPG-Festphase gebunden. Hierbei ist die 3'-Hydroxygruppe des Nukleosids über einen basenlabilen Ester als Linker an den Träger verknüpft. Dadurch kann das fertig synthetisierte Oligonukleotid durch basische Bedingungen vom Träger abgespalten werden. Im ersten Schritt der Synthese wird die temporäre, säurelabile DMTr-Schutzgruppe am 5'-Ende durch 3%-ige DCA abgespalten, um ein weiteres Nukleosid selektiv koppeln zu können. Hierdurch kann die freie 5'-Hydroxygruppe nukleophil am Phosphoratom des anderen Nukleosids angreifen, welches zuvor mit einem Tetrazolderivat aktiviert wurde. Die Kopplungseffizienz des C8-AAF-dG Phosphoramidits war etwas niedrig, sodass die Kopplungszeit auf 30 min (2 x 15 min) verlängert wurde. Dadurch konnte die Kopplungseffizienz auf ca. 60% gesteigert und somit die Ausbeute erhöht werden. Die Kopplungseffizienz wurde durch die photometrische Messung der DMTr-Kationen bei 500 nm bestimmt.



Schema 4: Reaktionszyklus der automatischen DNA-Festphasensynthese.

Da die Kopplungen nicht quantitativ ablaufen, werden in einem Cappingschritt die freien 5'-Hydroxygruppen acyliert. Ansonsten könnten diese im nächsten Reaktionszyklus erneut reagieren, wodurch verkürzte Oligonukleotide entstehen können. Zum Schluss wird das Phosphit durch Iod zum Phosphat oxidiert. Anschließend kann der Reaktionszyklus durch Abspaltung der DMTr-Gruppe erneut ablaufen, bis die gewünschte Länge des Oligonukleotids erreicht wird. Wie bereits oben beschrieben wurden *ultramild*-Konditionen nach *Gillet et al.* angewendet, um die Abspaltung der basenlabilen Acetylgruppe des Acetylaminofluorens während der Entschützung zu vermeiden.^[44] Dafür wurde Phenoxycyessigsäure anstatt Acetanhydrid als Capping-Reagenz verwendet. Weiterhin ist die exozyklische Aminfunktion der *ultramild* Phosphoramidite von dA und dG mit einer *iso*-Propyl-Phenoxyacetylgruppe geschützt. Auch die Abspaltung des Oligonukleotids von der Festphase und die Entschützung erfolgte durch *ultramild*-Bedingungen. Das dabei verwendete β -Mercaptoethanol verhindert die Oxidation des C8-AAF-dG Addukts.^[167]

Nach Synthese der Oligonukleotide wurden diese mittels präparativer RP-HPLC aufgereinigt. Dabei wurde ein Gradient von 30% Puffer B in Puffer A in 45 min gewählt. Für längere Oligonukleotide (ab 25mer) wurde ein längerer Gradient gewählt, um die Abbruchfragmente sauber und besser abzutrennen (25% Puffer B in Puffer A in 65 min). Durch diese beiden Reinigungsmethoden konnten alle Oligonukleotide mit einer Reinheit von 99% erhalten werden. Zusätzlich wurden die

Oligonukleotide *via* MALDI-TOF charakterisiert. Abbildung 20 stellt beispielhaft das präparative und das analytische HPL-Chromatogramm von einem C8-AAF-dG geschädigten 33mer dar.

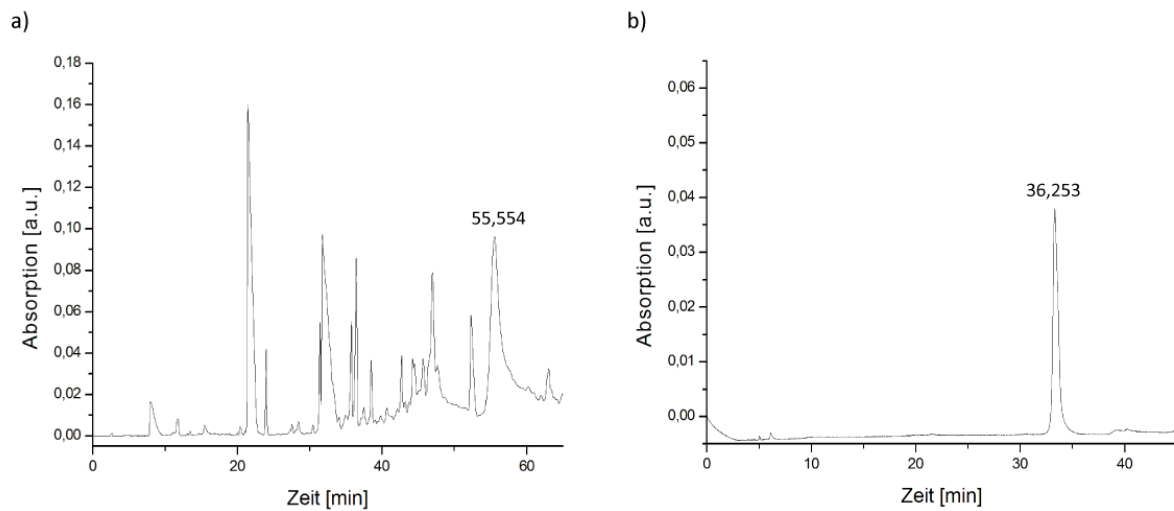


Abbildung 20: HPL-Chromatogramme von einem 33mer mit einem C8-AAF-dG Schaden. x-Achse: Retentionszeit [min]; y-Achse: Absorption bei 256 nm [a.u.]. **a) präparatives Chromatogramm;** Gradient 25% Puffer B in 65 min. **b) analytisches Chromatogramm nach der Aufreinigung;** Gradient: 30% Puffer B in 45 min.

4.3.2. Das humane XPA Protein

4.3.2.1. Allgemeines

Um weitere Einblicke in den NER-Mechanismus zu bekommen, sollte ein XPA/RPA/DNA-Komplex isoliert und untersucht werden. Dafür war die Expression und Reinigung des Volllängen Proteins XPA notwendig. Am Anfang wurde das Plasmid pPSG-IBA43-XPAfl von *Sandra Koch* verwendet. Zur gezielten Aufreinigung besitzt das XPA Protein ein C-terminalen Streptavidin-tagII und einen N-terminalen 6xHis-tag im pPSG-IBA43 Vektor. Die Aufreinigung ist in Abschnitt 3.1. beschrieben. Allerdings war die Ausbeute und Stabilität dieses Konstrukts nicht optimal, sodass ein neues XPA-Konstrukt mit nur einem tag, um die Stabilität des Proteins zu verbessern, kloniert wurde. Die Klonierung des neuen Konstrukts wird im folgenden Abschnitt beschrieben.

4.3.2.2. Klonierung

Die StarGate®-Technologie ermöglichte die Klonierung des *xpafI* Gens in den pPSG-IBA35 Vektor mit einem N-terminalen His-tag. Dabei wurde der Donorvektor IBA10 von *Sandra Koch* verwendet, wodurch nur noch die Transferreaktion in den pPSG-IBA35 Vektor durchgeführt werden musste. Die anschließende Transformation des Reaktionsansatzes in *E. coli* BL21(DE3)*pLysS* Zellen erwies sich als erfolgreich. Daneben wurde der Reaktionsansatz noch in *E. coli* NEB5α und BL21(DE3) Zellen transformiert. Allerdings war die Transferreaktion und Transformation in diese Zellen nicht erfolgreich. Die Überprüfung des Klonierungsprodukts erfolgte *via* Restriktionsverdau, welcher die gewünschten Fragmente im Agarosegel (Abbildung 21) zeigte. Zusätzlich wurde die richtige Sequenz des Vektors durch Sequenzierung bestätigt.

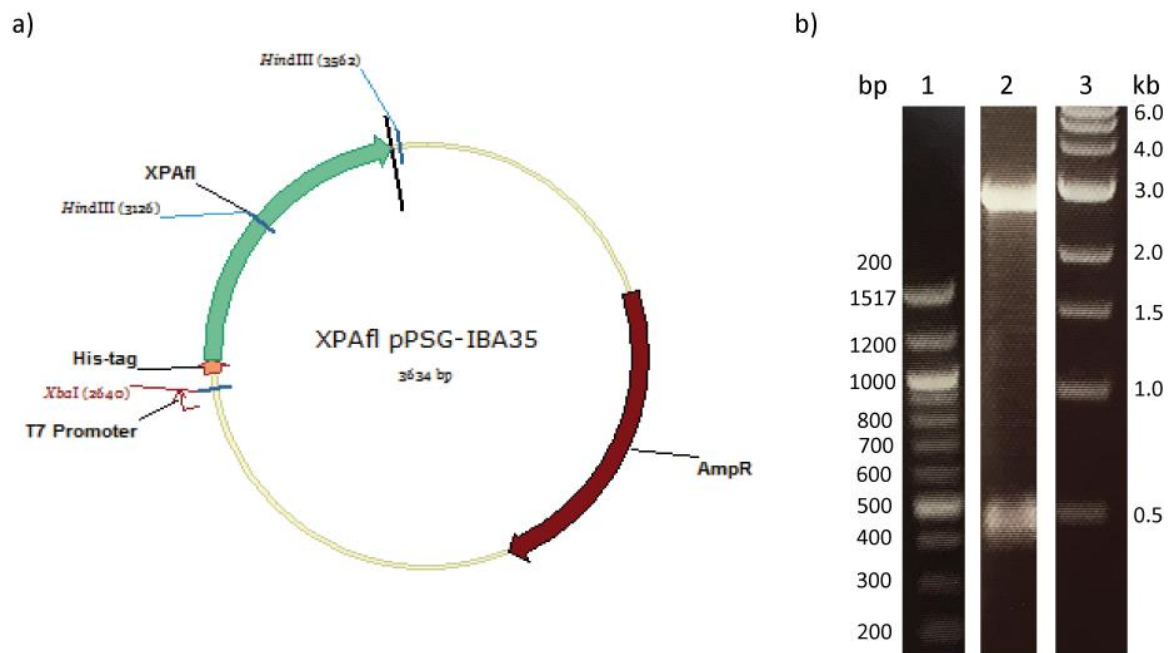


Abbildung 21: a) Vektorkarte von XPAfl pPSG-IBA35; rot: Ampicillin Resistenz; *HindIII* und *Xba*: Restriktionsstellen; orange: N-terminaler 6xHis-tag; grün: Protein; *HindIII* und *Xba*: Restriktionsstellen. b) 1%-iges Agarosegel des Restriktionsverdau von XPAfl pPSG-IBA35 mit den Restriktionsenzymen *HindIII* und *Xba*; Laufbedingungen: 100 V und 60 min; Spur 1: DNA Standard 100 bp; Spur 2: Restriktionsverdau mit den Fragmenten 2712, 483 und 460 bp; Spur 3: DNA Standard: 1 kb.

4.3.2.3. Heterologe Expression und Aufreinigung von XPA

Die heterologe Proteinexpression und Reinigung des zuvor hergestellten Klonierungsprodukts (XPAfl pPSG-IBA35) wurde durch verschiedene Bedingungen optimiert, wodurch die Proteinausbeute gesteigert werden konnte. Zu Beginn der Proteinexpression war die Menge an exprimiertem XPA nicht optimal, sodass weitere Expressionszellen wie BL21(DE3) oder Rosetta2(DE3) getestet wurden. Diese zeigten aber keinen signifikanten Unterschied in der Expression von XPA auf. Daneben wurden noch verschiedene Testexpressionen mit unterschiedlicher Induktionstemperatur, Induktionszeit, Konzentration an IPTG und Salzgehalt des LB-Mediums angesetzt. Dadurch konnten einige Parameter optimiert werden, die zu einer verbesserten Expression an XPA führten. Die Evaluierung der Testexpression erfolgte durch Probenentnahme zu bestimmten Zeitpunkten und anschließender Analyse *via* SDS-Gel. Tabelle 16 zeigt die Resultate, die zu einer Optimierung der Expression führten. Daneben konnte ein weiterer Parameter verbessert werden, indem die Übernachtskultur zuerst 1 h bei 37 °C und 200 rpm und danach 12 h bei 37 °C und 0 rpm inkubiert wurde anstatt 12 h bei 37 °C und 200 rpm.

Tabelle 16: Darstellung der optimierten Expressionsbedingungen von XPAfl pPSG-IBA35.

Protein XPA	Optimierte Bedingungen
<i>E. coli</i> Stamm	BL21(DE3) <i>pLysS</i>
LB-Medium	Mit 7.5 g/L NaCl
Übernacht-Kultur	500 mL 1h bei 220 rpm, dann 0 rpm über Nacht
OD ₆₀₀	0.6
Induktion	0.5 mM IPTG
Induktionstemperatur	37 °C
Induktionszeit	4 h
Schüttlergeschwindigkeit	200 rpm
Zusatz	10 µM ZnCl ₂

Außerdem konnte eine große Menge an XPA im Pellet nach dem Zellaufschluss festgestellt werden. Vermutlich durch die Bildung von Einschlusskörperchen (*inclusion bodies*) während der Expression oder durch einen unvollständigen Zellaufschluss. Daher wurde versucht, den Zellaufschluss *via* Ultraschall zu optimieren, indem verschiedene Anzahl an Zyklen, Zykluszeiten und Pulseinstellungen getestet wurden. Dabei konnte bei maximaler Energie, einer Pulseinstellung von 5 und 10 Zyklen für je 30 Sekunden das beste Ergebnis erhalten werden und die Menge an XPA im Pellet nach dem Zellaufschluss minimiert werden (Abbildung 22). Zusätzlich wurde ein Lysozym-Verdau vor der Ultraschall-Lyse durchgeführt, um den Zellaufschluss zu unterstützen.

Zum Schluss wurde die Reinigung des Proteins aus dem Zelllysät verbessert, um eine hohe Reinheit für die Kristallisationsstudien sowie für die Kryo-EM zu erreichen. Abbildung 22 zeigt die einzelnen Chromatogramme der entwickelten Reinigungsschritte und die Analyse der Reinigungseffizienz *via* SDS-PAGE. In der SDS-PAGE ist erkennbar, dass XPA ein Molekulargewicht von 40 kDa anstatt von 32.6 kDa besitzt (Abbildung 22e). Dieses Laufverhalten ist auf eine geladene Glutamin reiche Domäne im Protein sowie auf ungeordnete Bereiche im N- und C-Terminus zurückzuführen.^[168-170] Ein zusätzlich, eingebauter Chromatographieschritt (Abbildung 22b, Anionenaustauschchromatographie) im Reinigungsprotokoll ermöglichte eine Abtrennung von niedermolekularen Verunreinigungen. Weiterhin wurde jeder einzelner Chromatographieschritt optimiert. Dafür wurde im ersten Reinigungsschritt (Abbildung 22a, Ni-NTA-Affinitätschromatographie) die Anfangskonzentration von Imidazol von 20 mM auf 30 mM erhöht, um unspezifische Bedingungen zu vermindern. Zudem wurde der lineare Gradient zur XPA-Elution von 7 auf 8 Säulenvolumen erhöht. Dadurch konnten hochmolekulare Verunreinigungen, insbesondere eine Bande bei 70 kDa

abgetrennt werden. Dabei eluierte XPA bei 55% NiNTA-Puffer B. Des Weiteren wurde auch eine selbst beladene Zn-NTA-Säule anstatt einer Ni-NTA-Säule im ersten Chromatographieschritt getestet. Dabei sollte die Verwendung von Zink zu einer Stabilisierung des Zinkfingers von XPA führen, da Nickel den Zinkfinger beeinträchtigen könnte. Die Verwendung dieser Säule führte zu einer etwas besseren Reinheit nach dem ersten Chromatographieschritt. Allerdings konnte die selbst beladene Zn-NTA-Säule nur einmal benutzt werden, da diese nach jedem Gebrauch neu beladen werden musste. Folglich wurde die Ni-NTA-Säule für die Reinigung aus Zeitgründen weiter verwendet. Für den zweiten und dritten Chromatographieschritt (Abbildung 22b und c) wurde zusätzlich mit 5% Puffer B gewaschen, um weitere unspezifische Proteine zu eluieren. Auch hierbei wurde die Länge der Gradienten verlängert. Dabei wurden im zweiten Chromatographieschritt der Gradient von 13 auf 15 Säulenvolumen erhöht, wodurch insbesondere eine Verunreinigung bei 14 kDa und 35 kDa abgetrennt werden konnte (Abbildung 22e) und XPA bei 41% MonoQ-Puffer B eluierte. Vermutlich sind diese Verunreinigungen Abbauprodukte von XPA. Die Erhöhung des Gradienten von 30 auf 35 Säulenvolumen während der Heparinsäule (Abbildung 22c) ermöglichte eine bessere Abtrennung von nieder- als auch von hochmolekularen Verunreinigungen. Hinzu konnte XPA bei 50% Heparin-Puffer B mit einer Reinheit von 90% eluiert werden.

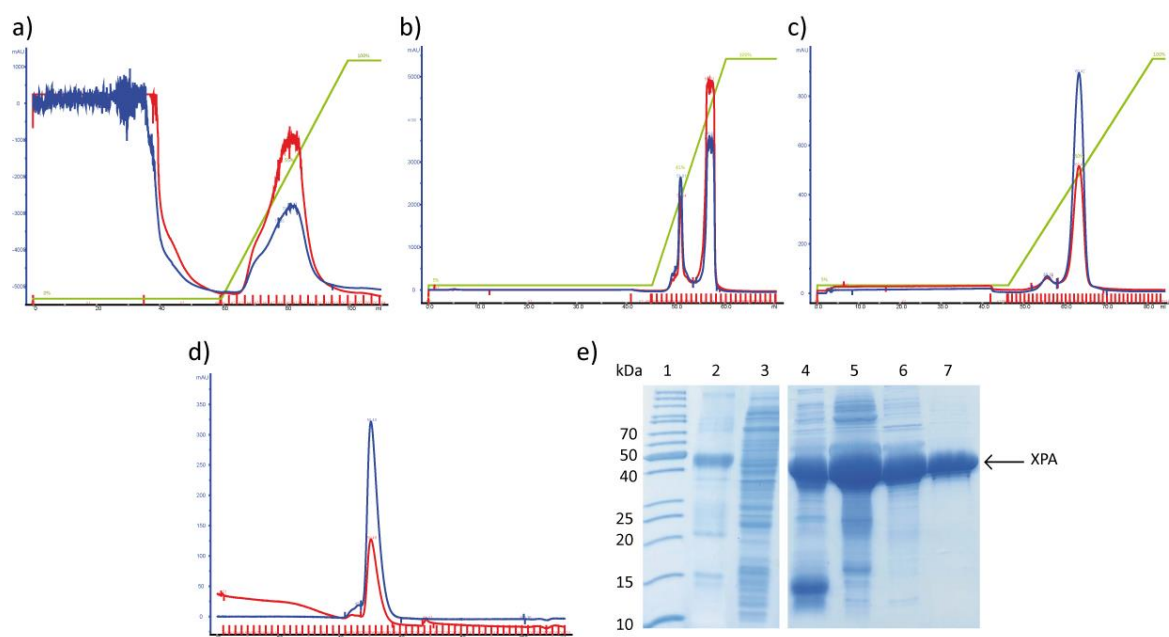


Abbildung 22: Chromatogramme und SDS-PAGE der Reinigungsschritte von XPA. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. a) Ni-NTA Affinitätschromatographie. b) MonoQ Anionenaustauschchromatographie. c) Heparin Affinitätschromatographie. d) Gelfiltrationschromatographie. e) Coomassie gefärbtes 15%iges SDS-PAGE. Laufbedingungen: 180 V, 1h. 1: Proteinstandard PageRuler™, 2: Pellet nach dem Aufschluss; 3: Zellysat; 4: Eluat der Ni-NTA-Säule bei 55% Puffer B; 5: Eluat der MonoQ-Säule bei 41% Puffer B; 6: Eluat der Heparin-Säule bei 50% Puffer B; 7: Eluat der Gelfiltrationssäule bei einem Retentionsvolumen von 50.45 mL.

Der letzte Chromatographieschritt sollte letzte Verunreinigungen und Aggregate abtrennen. Dabei konnte XPA mit einer Reinheit von 99% (Abbildung 22e, Spur 7) und mit einem Retentionsvolumen von 55.45 mL erhalten werden. Ferner betrug die Proteinausbeute 3-4 mg aus 10 L Kultur und konnte durch die oben beschriebenen Parameter signifikant verbessert werden (zuvor ca. 1 mg aus 10 L und 90% rein). Im Vergleich dazu betrug die Ausbeute von XPA pPSG-IBA43 nur 1 mg aus 10 L. Auch die Stabilität von XPA konnte verbessert werden, da weniger Aggregate im Vergleich zu XPA pPSG-IBA43 während der Gelfiltration (Abbildung 22d) erkennbar waren. Auch die Zugabe von 5 mM anstatt von 1 mM DTT zu den einzelnen Puffern steigerte die Stabilität von XPA. Dies äußerte sich vor allem in der geringeren Aggregatbildung des Proteins.

4.3.3. Das humane Protein RPA

4.3.3.1. Allgemeines

Für die Untersuchung eines XPA/RPA/DNA-Komplexes ist die Expression und Aufreinigung von RPA mit allen drei Untereinheiten notwendig. Das Plasmid p11d-RPA wurde freundlicherweise von Prof. Marc Wold (Universität Iowa, USA) bereitgestellt. Das Plasmid besitzt alle drei Untereinheiten von RPA und keine Sequenz zur Epitopmarkierung des Proteins. Die Expression und Aufreinigung erfolgte nach dem Protokoll von *Wold et al.*^[156]

4.3.3.2. Aufreinigung von RPA

RPA ist für *E. coli* giftig, sodass für jede Expression eine neue Transformation erfolgte. Zudem konnte eine stärkere Expression von RPA32 im Gegensatz zu den anderen beiden Untereinheiten detektiert werden, welches bereits von *Wold et al.* beschrieben werden konnte.^[156] Für die anschließende Zelllyse wurde das entwickelte Protokoll *via* Ultraschall-Lyse von XPA verwendet.

Abbildung 23 veranschaulicht die einzelnen Chromatogramme der Reinigungsschritte und die Analyse der Reinigungseffizienz *via* SDS-PAGE. Die Zugabe von DTT zu den einzelnen Puffern kurz vor der Aufreinigung sollte die Stabilität von RPA erhöhen, da dieses vor unerwünschter Oxidation von Cysteinresten schützt. Im ersten Reinigungsschritt (Abbildung 23a) konnte RPA mit einer Konzentration von 1.5 M NaSCN (100% HI-Puffer D) von der GEBLue-Affinitätssäule eluiert werden. Denn NaSCN ist ein chaotropes Salz, welches die Dissoziation von RPA von der Matrix verursacht. Die zuvor durchgeführten Waschschrte (80% HI-B und 100% HI-C) ermöglichten die Abtrennung von vielen unspezifischen Proteinen, sodass eine Reinheit des RPA Proteins von ca. 60% erreicht werden konnte (Abbildung 23d, Spur 4). Außerdem führte das NaSCN-Salz in der Proteinprobe zu einem Schmier im SDS-Gel. Dieser Schmier konnte durch Acetonfällung der SDS-Probe und Aufnahme des Proteinpellets in Wasser vermindert werden (Abbildung 23d, Spur 4). Im nächsten Reinigungsschritt wurde eine HAP-Affinitätschromatographie anstatt einer Dialyse durchgeführt, um das NaSCN-Salz zu entfernen (Abbildung 23b). Da das NaSCN-Salz RPA teilweise reversibel denaturieren kann, wurde der zweite Chromatographieschritt sehr zügig nach dem ersten Reinigungsschritt ausgeführt. Dabei konnte RPA mit 80 mM KH₂PO₄ (100% HI-Puffer F) und einer Reinheit von 90% eluiert werden (Abbildung 22d, Spur 5). Neben der Entfernung des NaSCN Salzes konnten viele nieder- und hochmolekulare Verunreinigungen entfernt werden. Das Eluat der HAP-Affinitätschromatographie wurde für den letzten Reinigungsschritt mit HI-Puffer A verdünnt, um die ionische Stärke im Eluat zu reduzieren. Der anschließende Reinigungsschritt *via* Anionenaustauscher (Abbildung 23c)

ermöglichte durch einen Stufengradienten die Entfernung von restlichen Verunreinigungen, insbesondere die Bande bei ca. 40 kDa. Dabei konnte RPA mit 350 mM KCl (35% HI-Puffer B) und einer Reinheit von ca. 99% eluiert werden (Abbildung 23d, Spur 6). Für weitere Versuche wurde das RPA Protein im Kristallisationspuffer II umgepuffert.

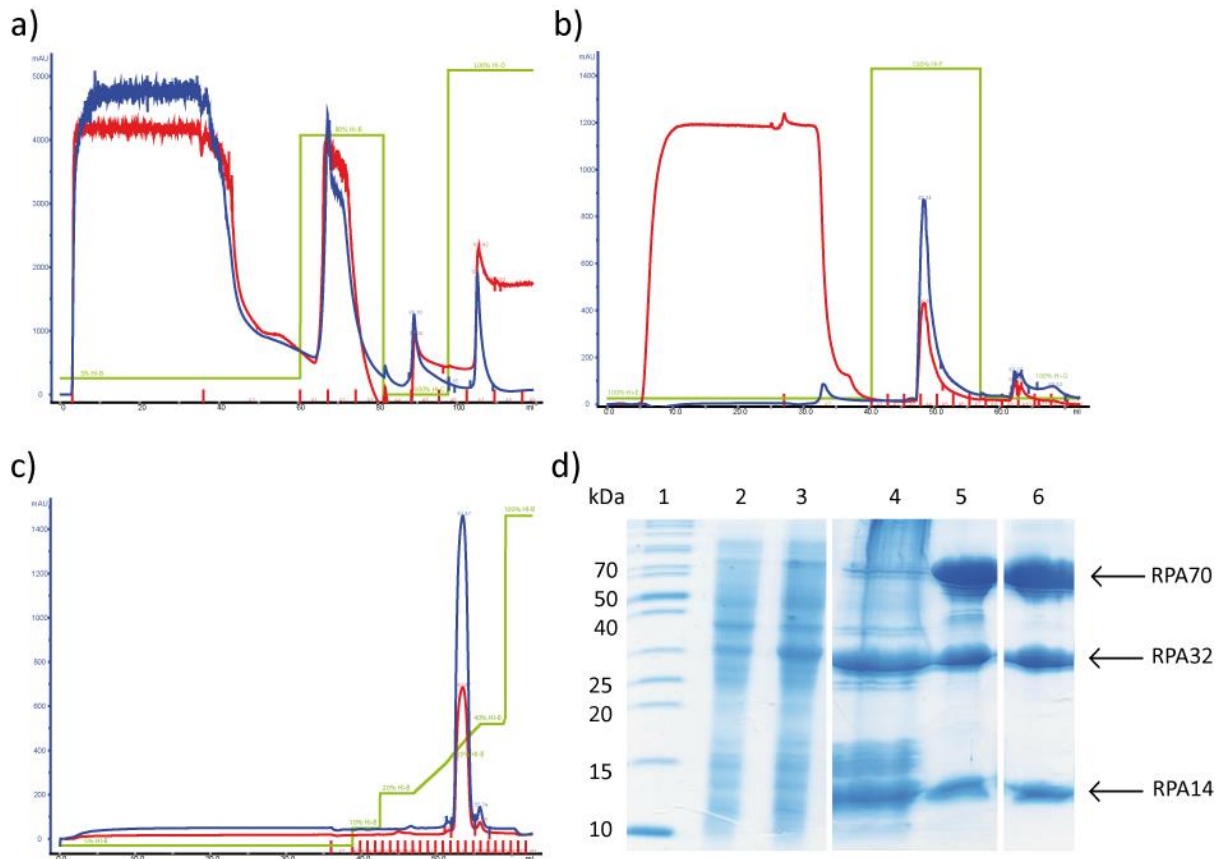


Abbildung 23: Chromatogramme und SDS-PAGE der Reinigungsschritte von RPA. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. a) GELBlue-Affinitätschromatographie. b) HAP-Affinitätschromatographie. c) MonoQ-Anionenaustauschchromatographie. d) Coomassie gefärbtes 15%iges SDS-PAGE. Laufbedingungen: 180 V, 1h. 1: Proteinstandard PageRuler™, 2: Pellet nach dem Aufschluss; 3: Zelllysatz; 4: Eluat der GELBlue-Säule bei 100% HI-Puffer D; 5: Eluat der HAP-Säule bei 100% HI-Puffer F; 6: Eluat der MonoQ-Säule bei 35% HI-Puffer B.

Weiterhin zeigt das SDS-Gel, dass ein 1:1:1 Komplex der drei Untereinheiten von RPA aufgereinigt werden konnte. Denn die Banden-Intensitäten im SDS-Gel nehmen proportional zum Molekulargewicht ab. Ferner betrug die Proteinausbeute 4.5-6.0 mg aus 5 L Kultur, welche mit den Literaturwerten von 0.5-1.0 mg/L übereinstimmt.^[156]

4.3.4. Isolierung eines stabilen XPA/RPA/DNA-Komplexes

4.3.4.1. Allgemeines

Nach Aufreinigung der Proteine XPA und RPA, sollten optimale Bedingungen für strukturelle Studien eines stabilen XPA/RPA/DNA-Komplexes gefunden werden. Titrationsstudien von *Yang et al.* zeigten bereits, dass sich ein XPA₂/RPA-Komplex ausbildet, welcher mittels einer Gelfiltrationssäule isoliert werden konnte.^[137] Basierend auf diesen Studien sollte ein Gelfiltrationsassay entwickelt werden, der den XPA/RPA/DNA-Komplex isoliert und charakterisiert. Zudem sollten mehrere DNA-Substrate getestet werden, um eine geeignete Sequenz für die Komplexbildung zu finden. Zu Beginn der Entwicklung des Assays wurden die Proteine im Kristallisationspuffer I gelagert, welcher 5% Glycerin enthält. Für spätere Strukturstudien, wie Kryo-EM, ist ein zu hoher Anteil an Glycerin nicht empfehlenswert, da dies die Viskosität der Probe beim Einfrieren erhöht. Daher wurden noch verschiedene Pufferbedingungen für beide Proteine getestet, die sowohl weniger Glycerin enthalten als auch die Stabilität der Proteine nicht beeinträchtigen.

4.3.4.2. Thermofluor Assay

Mit Hilfe des Thermofluor Assays wurden verschiedene Pufferbedingungen in einem ersten Screen nach *Boivin et al.* für die Proteine XPA und RPA getestet.^[162] Dabei wurden unterschiedliche Salze, pH-Werte und Additive getestet, um die Stabilität der Proteine zu untersuchen. Daneben wurde in einem zweiten Screen unterschiedliche Mengen an Glycerin im Kristallisationspuffer I getestet, um eine zusätzliche Anwendung in Kryo-EM Studien zu ermöglichen (die verschiedenen Screenings befinden sich im Anhang 6.4.). Abbildung 24 zeigt die Ergebnisse des Thermofluor Assays (Pufferbedingungen nach *Boivin et al.*^[162]) mit den am besten gemeinsamen Pufferbedingungen für XPA und RPA zusammen. Dabei wurde die Fluoreszenz des Farbstoffs *SYPRO Orange* in Abhängigkeit der Temperatur gemessen. Die genaue Funktionsweise des Thermofluor Assays wurde bereits im Abschnitt 4.2.2.5.3. beschrieben.

Dabei ist in den Evaluierungsergebnissen (Abbildung 24) ein signifikanter Unterschied in der Fluoreszenz zwischen RPA und XPA erkennbar. Der Thermofluor Assay weist für XPA ein geringes Fluoreszenzsignal auf. Zunächst wurde angenommen, dass dies auf eine sehr geringe Stabilität von XPA in den Bedingungen nach *Boivin et al.* hinweist, wodurch XPA in den jeweiligen Puffern denaturiert vorliegen könnte und der Farbstoff somit bereits vor der Messung an XPA bindet.^[162] Da XPA im Kristallisationspuffer I stabil vorlag, wurden, wie bereits oben erwähnt, verschiedene Glycerinmengen im Kristallisationspuffer I getestet. Weiterhin wurden dieselben Bedingungen für

beide Proteine zusammen untersucht. Dabei wurde angenommen, dass sich die Proteine gegenseitig stabilisieren und dadurch die Identifizierung guter Pufferbedingungen für XPA möglich wäre. Allerdings führten beide Ansätze zu kaum einer Fluoreszenz und somit zu keinem Erfolg. Wahrscheinlich kann XPA nicht über einen Thermofluor Assay analysiert werden, da XPA viele hydrophobe Bereiche besitzt und der Farbstoff vermutlich vor der Messung an XPA bindet. Des Weiteren präzipitierte XPA sichtbar bei einer Verringerung des Glyceringehalts. Deswegen wurde zunächst der Kristallisationspuffer I mit 5% Glycerin für die Komplexisolierung verwendet. Zum einen ist XPA in diesem Puffer stabil und zum anderen sind 5% Glycerin gerade noch am Grenzwert für Kryo-EM Bedingungen. Im Gegensatz dazu wies RPA in vielen verschiedenen Pufferbedingungen (Abbildung 24) eine hohe Stabilität auf, so auch im Kristallisationspuffer I.

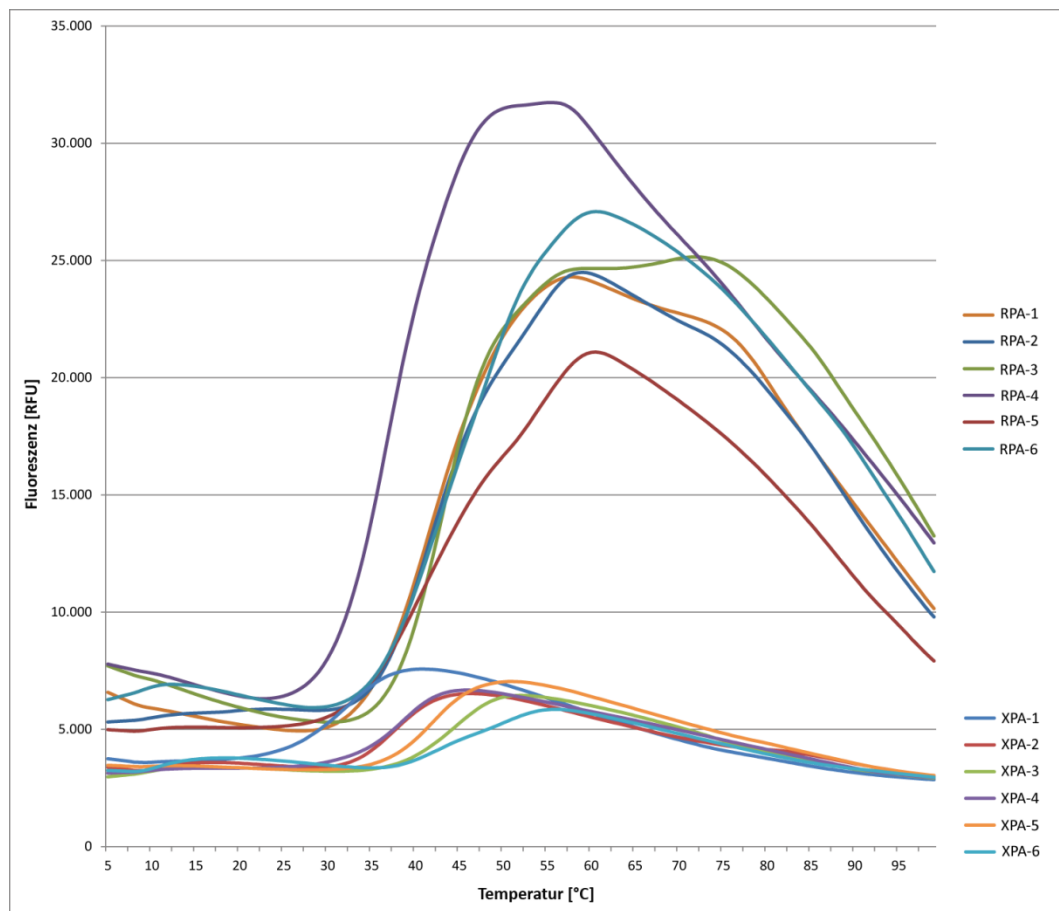


Abbildung 24: Ergebnis des Thermofluor Assays von XPA und RPA. x-Achse: Temperatur [°C]; y-Achse: Fluoreszenz [RFU]; rechts: Legende der verwendeten Puffern für RPA und XPA 1: Natriumkakodylat pH 6.5; 2: HEPES pH 7.0; 3: Tris-HCl pH 7.5; 4: Imidazol pH 8.0; 5: Tricin pH 8.0; 6: Tris-HCl pH 8.5.

4.3.4.3. Gelfiltrationsassay

4.3.4.3.1. Gelfiltration eines Standards und der einzelnen Proteine

Zur Isolierung und zum Nachweis eines XPA/RPA/DNA-Komplexes sollte ein Gelfiltrationsassay entwickelt werden. Die dabei verwendete, analytische Gelfiltrationssäule (Superdex™ 200 Increase 10/300 GL, 24 mL), beruht auf dem SEC-Prinzip, wodurch der Komplex im Gegensatz zu den einzelnen Komponenten früher eluieren sollte, da dieser im Vergleich zu den einzelnen Proteinen und der DNA größer ist. *Yang et al.* nahmen an, dass XPA als Dimer vorliegt.^[137] Um das Molekulargewicht der, in dieser Arbeit verwendeten, Proteine XPA und RPA bestimmen zu können, wurde ein Gelfiltrationsstandard von Biorad, welcher die Proteine Thyroglobulin, γ -Globulin, Ovalbumin, Myoglobin und Vitamin B12 enthält, verwendet. Abbildung 25 zeigt die Chromatogramme des Standards und der Einzelproteine. XPA und RPA enthaltende Fraktionen wurden durch die Absorption bei 280 nm identifiziert und mittels SDS-PAGE analysiert (Abbildung 25d).

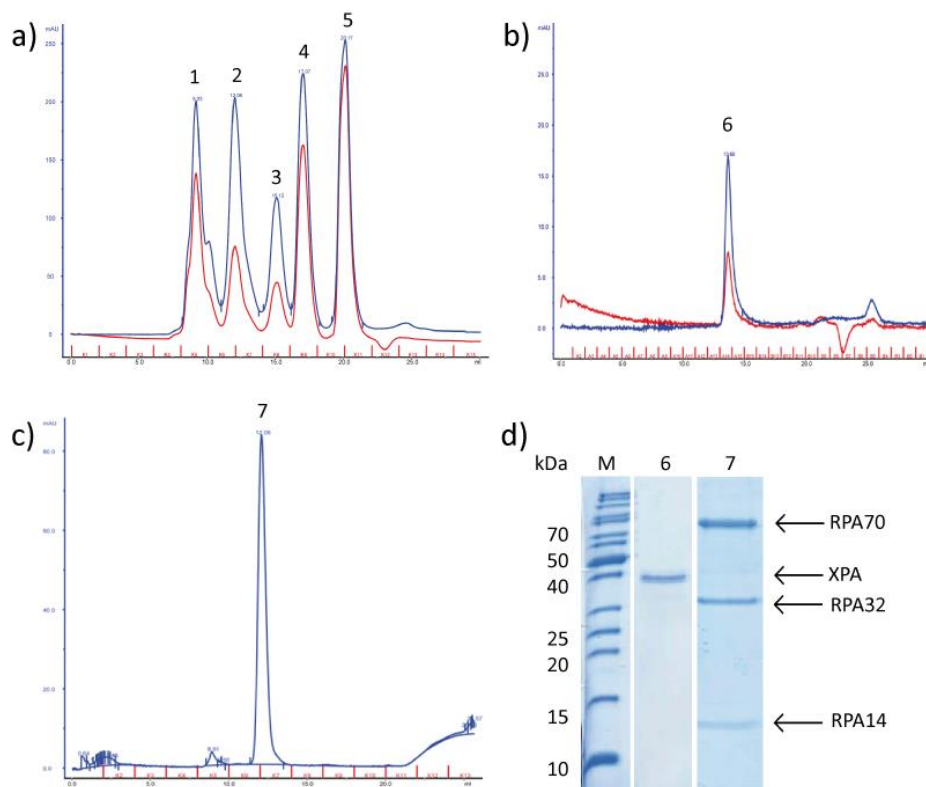


Abbildung 25: Chromatogramme des Standards und der Einzelproteine sowie SDS-PAGE zur Identifizierung. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL). x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. a) Gelfiltration des Standards; Retentionsvolumen von: Thyroglobulin 9.23 mL (1), γ -Globulin 12.06 mL (2), Ovalbumin 15.08 mL (3), Myoglobin 17.04 mL (4) und Vitamin B12 20.12 mL (5). b) Gelfiltration von XPA; Retentionsvolumen von XPA 13.68 mL (6). c) Gelfiltration von RPA; Retentionsvolumen von RPA 12.08 mL (7). d) Coomassie gefärbtes 15%iges SDS-Gel. Laufbedingungen: 180 V, 1h. M: Proteinstandard PageRuler™; 6: Eluat der Gelfiltration von XPA; 7: Eluat der Gelfiltration von RPA.

Zur Evaluierung der molekularen Masse von XPA und RPA wurde der Logarithmus aus den molekularen Massen der Proteine vom Standard gebildet. Zudem wurden die Retentionsvolumina des Standards mehrfach über die Gelfiltrationssäule bestimmt, um einen Durchschnittswert für die Retentionsvolumina zu erhalten. Danach konnte durch den gebildeten Logarithmus aus den molekularen Massen des Standards in Abhängigkeit der Retentionsvolumina eine Kalibrationsgerade ($y = -4.121x + 33.699$) mit dem Programm Excel (Microsoft) berechnet werden. Durch die erhaltenen Retentionsvolumina von XPA und RPA kann die molekulare Masse der Proteine mit Hilfe der Kalibrationsgeraden bestimmt werden. Tabelle 17 zeigt die Retentionsvolumina, sowie die errechneten molekularen Massen.

Tabelle 17: Retentionsvolumen und molekulare Masse der einzelnen Proteine. Verwendete Kalibrationsgerade: $y = 4.121x + 33.699$.

Proteine	Retentionsvolumen [mL]	Molekulare Masse [kDa]	Berechnete molekulare Masse [kDa]
Thyroglobulin	9.23	670	
γ -Globulin	12.06	158	
Ovalbumin	15.08	44	
Myoglobin	17.04	17	
Vitamin B12	20.12	1.35	
XPA	13.68	32.62	72
RPA	12.08	110	176

Die mittels Gelfiltrationsstandard berechneten molekularen Massen für XPA₂ (72 kDa) und RPA (176 kDa) stimmen mit den Werten von *Yang et al.* überein.^[137] Den etwas zu hohen Wert von RPA (176 kDa anstatt 110 kDa) erklären diese durch die nicht-globuläre Konformation von RPA, wodurch sich das Laufverhalten von RPA auf einer Gelfiltrationssäule ändert. Daneben vermuten *Yang et al.*, dass XPA in ihrer Titrationsstudie als Homodimer vorliegt.^[137] Dieses würde auch die zu hohe molekulare Masse von XPA (72 kDa anstatt 32.62 kDa) in dieser Arbeit erklären. Die erhaltenen Ergebnisse belegen die Daten von *Yang et al.*^[137] und könnten Hinweise auf die Zusammensetzung des XPA/RPA/DNA-Komplexes geben.

4.3.4.3.2. Isolierung des XPA/RPA/DNA-Komplexes durch einen Gelfiltrationsassay

Die Idee des Gelfiltrationsassays bestand darin, den XPA/RPA/DNA-Komplex über die Gelfiltrationssäule zu isolieren und anschließend mittels SDS-PAGE und nativen Gel zu charakterisieren. Da RPA eine längere DNA (20-30nt) stabiler binden kann, wurde als erstes ein 34mer als DNA-Substrat für die Komplexbildung ausgewählt und verwendet.^[57] Hierbei wurde XPA zusammen mit RPA und doppelsträngiger AAF-geschädigter DNA (34mer_1 und 34mer_3) in einem Verhältnis von 4:4:1 bei 4 °C für 30 min inkubiert. Nach anschließender Zentrifugation wurde der Überstand auf die Gelfiltrationssäule aufgetragen. Durch die Absorption bei 260 nm und 280 nm wurden Protein und DNA enthaltende Fraktionen identifiziert. Diese wurden zum einen mittels SDS-PAGE analysiert, um den XPA/RPA/DNA-Komplex enthaltende Fraktionen zu identifizieren. Zum anderen wurden die Fraktionen mit einem nativen Gel analysiert, um DNA enthaltende Fraktionen zu identifizieren und Protein-DNA Interaktionen nachweisen zu können. Abbildung 26 zeigt die ersten Ergebnisse des Gelfiltrationsassays.

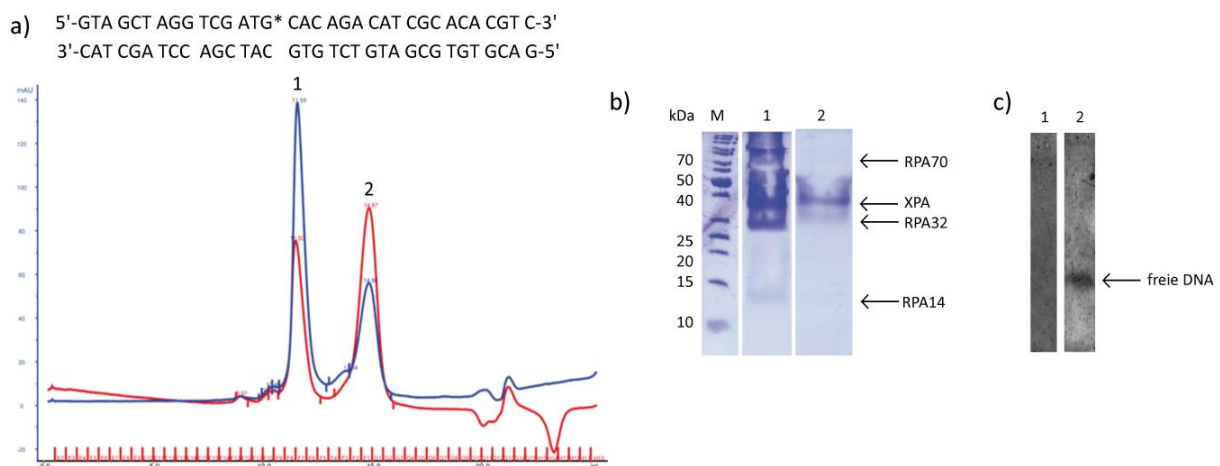


Abbildung 26: a) Chromatogramm des Gelfiltrationsassays mit der XPA/RPA/DNA-Probe im Verhältnis von 4:4:1 und die Sequenz der verwendeten DNA. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL). x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. Die beiden Hauptpeaks haben ein Retentionsvolumen von 11.59 mL (1) und 14.67 mL (2). b) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltenden Fraktionen. M: Proteinstandard PageRuler™, 1: Eluat bei 11.59 mL; 2: Eluat bei 14.67 mL. c) SYBR Green I gefärbtes natives Gel der DNA enthaltenden Fraktionen. 1: Eluat bei 11.59 mL; 2: Eluat bei 14.67 mL.

Bei der Gelfiltration konnten (Abbildung 26a) zwei Peaks bei einem Retentionsvolumen von 11.59 mL und 14.67 mL erhalten werden. Bei Betrachtung der 260 nm und 280 nm Absorption ist zu erkennen, dass das Verhältnis von 280:260 nm 2:1 beim ersten Peak und beim zweiten Peak 1:2 ist. Dies könnte bedeuten, dass der zweite Peak hauptsächlich DNA enthält. Nach Analyse beider Peaks mittels SDS-PAGE (Abbildung 26b) konnten die Proteine beider Eluate identifiziert werden. Der erste Peak enthielt einen XPA/RPA-Komplex und der zweite Peak nur XPA. Die Fraktionen wurden als nächstes über ein natives Gel charakterisiert, um DNA enthaltende Eluate zu identifizieren. Dabei

konnte keine DNA im ersten Peak detektiert werden, wodurch ein XPA/RPA-Komplex in dieser Fraktion vorliegen sollte. Die Analyse der Bandenintensität des XPA/RPA-Komplex mittels SDS-PAGE zeigte zudem, dass ein XPA₂/RPA-Komplex vorliegt. Diese Beobachtungen stimmen mit den Ergebnissen der XPA/RPA-Titrationsstudie von *Yang et al.* überein.^[137] Im Gegensatz dazu konnte DNA im zweiten Peak detektiert werden. Somit könnte der zweite Peak vielleicht einen XPA/DNA-Komplex enthalten. Allerdings ist kein DNA-Protein-Shift im nativen Gel erkennbar, sodass der XPA/DNA-Komplex wahrscheinlich nur in geringen Mengen vorlag und hauptsächlich freie DNA in dieser Fraktion vorhanden war. Letzten Endes zeigten diese ersten Ergebnisse, dass sich kein XPA/RPA/DNA-Komplex, sondern ein XPA/RPA-Komplex ausbildet. Eine weitere Optimierung des Gelfiltrationsassays sollte unternommen werden, um eine Ausbildung des XPA/RPA/DNA-Komplexes zu erzielen. Daneben wurden XPA zusammen mit RPA, sowie die DNA einzeln über die Gelfiltrationssäule analysiert, um die einzelnen Retentionsvolumina bestimmen zu können. Ein XPA/RPA/DNA-Komplex müsste ein geringeres Retentionsvolumen im Vergleich zu den einzelnen Komponenten haben. Durch die Bestimmung einzelner Retentionsvolumina können die Peaks im Chromatogramm des Gelfiltrationsassays leichter identifiziert werden. Tabelle 18 zeigt die Retentionsvolumina der einzelnen Komponenten sowie deren molekulare Masse.

Tabelle 18: Retentionsvolumen und molekulare Masse der einzelnen Proteine. Verwendete Kalibrationsgerade aus Abschnitt 4.3.4.3.1: $y = 4.121x + 33.699$.

Proteine	Retentionsvolumen [mL]	Molekulare Masse [kDa]	Berechnete molekulare Masse [kDa]
XPA	13.68	32.62	72
RPA	12.08	110	176
ds AAF(C8)-dG 34mer	14.68	10.65	-
XPA/RPA (im Verhältnis 2:1 inkubiert)	11.59	175.24 (XPA:RPA2:1)	232

Die Retentionsvolumina von DNA und XPA/RPA (Tabelle 18) stimmen mit den Werten aus Abbildung 26 überein. Die molekulare Masse konnte mit Hilfe der Kalibrationsgeraden aus Abschnitt 4.3.4.3.1. berechnet werden. Demnach wurde ein XPA₂/RPA-Komplex und ein DNA-Peak erhalten.

4.3.4.3.3. Optimierung des Gelfiltrationsassays

Nach den ersten Studien des Gelfiltrationsassays sollte dieser optimiert werden, sodass ein XPA/RPA/DNA-Komplex isoliert werden kann. Dabei sollte unter anderem der verwendete Puffer verändert werden. *Chen et al. zeigten*, dass RPA bei einer höheren Natriumchlorid Konzentration eine höhere Bindungsaffinität zu DNA hat.^[66] Ferner zeigten Fluoreszenzpolarisationsstudien von XPA und RPA, dass eine zu hohe Magnesiumkonzentration die Bindungsaffinität zu DNA ebenfalls reduziert.^[171] Basierend auf diesen Studien wurde die Magnesiumchlorid Konzentration im Kristallisationspuffer II reduziert und zusätzlich Natriumchlorid hinzugefügt. Tabelle 19 zeigt nochmal den direkten Vergleich beider Puffer. Weiterhin ist eine geringere Magnesiumchlorid Konzentration für die Kristallisationsstudien von Vorteil. Die Optimierung des Puffers erfolgte in Kooperation mit der Kisker Gruppe (Würzburg).

Tabelle 19: Vergleich des Kristallisationspuffers I und II.

Kristallisationspuffer I	Kristallisationspuffer II
50 mM HEPES-NaOH pH 8.0	50 mM HEPES-NaOH pH 8.0
5% (v/v) Glycerin	5% (v/v) Glycerin
50 mM MgCl₂·6H₂O	5 mM MgCl₂·6H₂O, 150 mM NaCl
5 mM DTT	5 mM DTT
10 µM ZnCl ₂	10 µM ZnCl ₂

Anschließend wurde XPA zusammen mit RPA und doppelsträngiger AAF-geschädigter DNA (34mer_1 und 34mer_3) im Kristallisationspuffer II in einem Verhältnis von 4:4:1 bei 4 °C für 30 min inkubiert. Danach wurde die Probe mittels Gelfiltration charakterisiert. Abbildung 27 zeigt einen Vergleich des Gelfiltrationsassays mit Kristallisationspuffer I (27a) und II (27b). Zudem wurden die Protein und DNA enthaltenden Fraktionen der Gelfiltration mit Kristallisationspuffer II *via* SDS-PAGE und nativen Gel analysiert (Abbildung 27c und d). Bei Vergleich beider Chromatogramme konnten bei beiden Puffern zwei Peaks mit fast identischen Retentionsvolumina erhalten werden (Kristallisationspuffer I: 11.59 mL und 14.67 mL, sowie Kristallisationspuffer II: 11.61 mL und 14.24 mL). Bei genauer Betrachtung der 260 nm Absorption mit Kristallisationspuffer II fällt auf, dass diese im ersten Peak höher ist als im zweiten (Abbildung 27b). Im Gegensatz dazu, ist die 260 nm Absorption mit Kristallisationspuffer I im zweiten Peak höher (Abbildung 27a). Dies lässt vermuten, dass der Peak bei 11.61 mL DNA enthalten könnte.

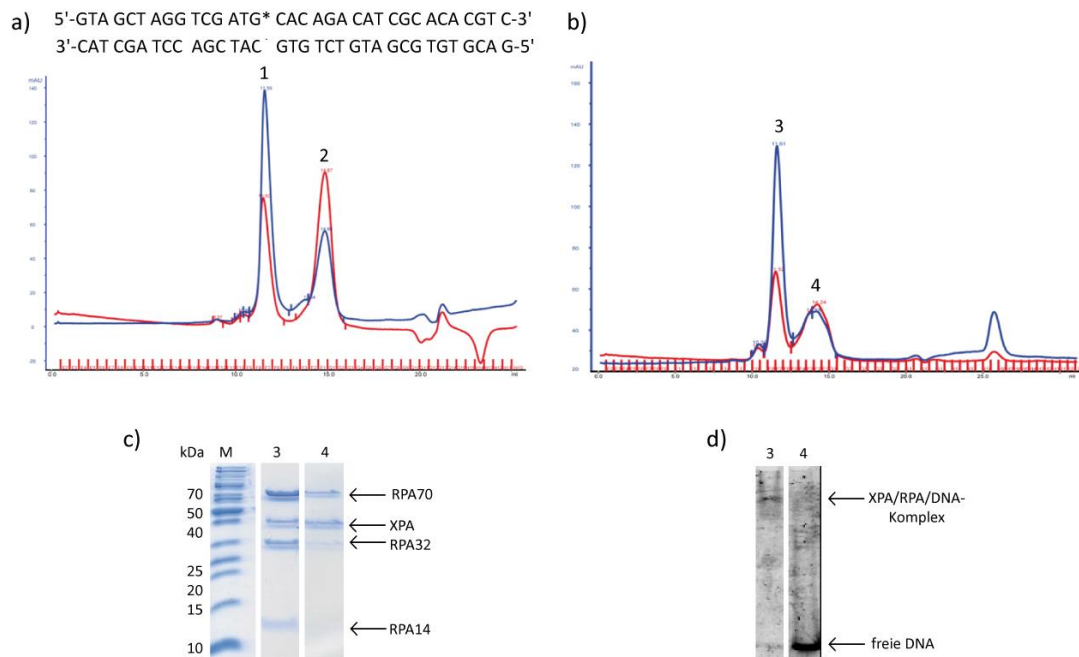


Abbildung 27: Vergleich der Chromatogramme von der XPA/RPA/DNA-Probe im Verhältnis von 4:4:1 mit a) Kristallisationspuffer I; mit b) Kristallisationspuffer II. Sequenz der verwendeten DNA mit G*: AAF(C8)-dG. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL). x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. Die Hauptpeaks beider Chromatogramme haben ein Retentionsvolumen von 11.59 mL (1), 14.67 mL (2), 11.61 mL (3) und 14.24 mL (4). **c) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltenden Fraktionen von 3 und 4.** M: Proteinstandard PageRuler™, 3: Eluat bei 11.61 mL; 4: Eluat bei 14.24 mL. **d) SYBR Green I gefärbtes natives Gel der DNA enthaltenden Fraktionen von 3 und 4.** 3: Eluat bei 11.61 mL; 4: Eluat bei 14.24 mL.

Nach Analyse beider Peaks mit Kristallisationspuffer II mittels SDS-PAGE (Abbildung 27c) konnten die Proteine beider Eluate identifiziert werden. Nach Abschätzung der Bandenintensität enthielt der erste Peak interessanterweise einen XPA/RPA-Komplex. Der zweite Peak enthielt hauptsächlich XPA mit geringen Mengen an RPA. Im Vergleich dazu enthielt der zweite Peak mit Kristallisationspuffer I (Abbildung 26b) nur XPA. Dies erklärt die leichte Verschiebung des Retentionsvolumens von 14.67 mL auf 14.24 mL. Die anschließende Charakterisierung der Fraktionen im Kristallisationspuffer II mit einem nativen Gel (Abbildung 27d) zeigten, dass im zweiten Peak mehr DNA vorlag. Dennoch enthielt der erste Peak diesmal etwas DNA, was auch durch einen DNA-Protein-Shift im nativen Gel belegt werden konnte. Der XPA/RPA-Komplex hatte somit etwas DNA gebunden. Zudem bildet sich aus dem XPA₂/RPA-Komplex ein XPA/RPA-Komplex, sobald dieser DNA bindet. Der zweite Peak wies hauptsächlich wieder freie ungebundene DNA auf.

Folglich ermöglichte der neue Puffer eine geringe DNA-Bindung der XPA/RPA-Proteine, dennoch war die DNA größtenteils immer noch ungebunden. Daher sollte als nächstes die DNA-Sequenz für den XPA/RPA/DNA-Komplex optimiert werden, wobei nun der neue Puffer weiterverwendet werden sollte. EMSA Studien von *Yang et al.* belegten, dass XPA und RPA zusammen an Y-förmige DNA binden, wobei XPA an der Verzweigung und RPA an die ss Bereiche der DNA bindet, wodurch die Interaktion von RPA mit XPA erleichtert wird.^[148] Hinzu konnten sie durch eine Gelfiltration zeigen,

dass XPA als Monomer an Y-förmige DNA im Vergleich zu nur doppelsträngiger (geschädigter) DNA bindet.^[148] Basierend auf dieser Studie sollte ein AAF-geschädigtes 34mer mit einem 20nt 5'-Überhang für die Komplexbildung getestet werden. Daraufhin wurde diese DNA (34mer_1 und 34mer_4) zusammen mit XPA und RPA in einem Verhältnis von 4:4:1 bei 4 °C für 30 min inkubiert. Danach wurde die Probe mittels Gelfiltration charakterisiert. Abbildung 28 zeigt den direkten Vergleich des Gelfiltrationsassays mit einem AAF-geschädigten 34mer ohne (28a) und mit einem 20nt 5'-Überhang (28b). Zudem wurden die Protein und DNA enthaltenden Fraktionen der Gelfiltration mit dem Überhang *via* SDS-PAGE und nativen Gel analysiert (Abbildung 28c und d).

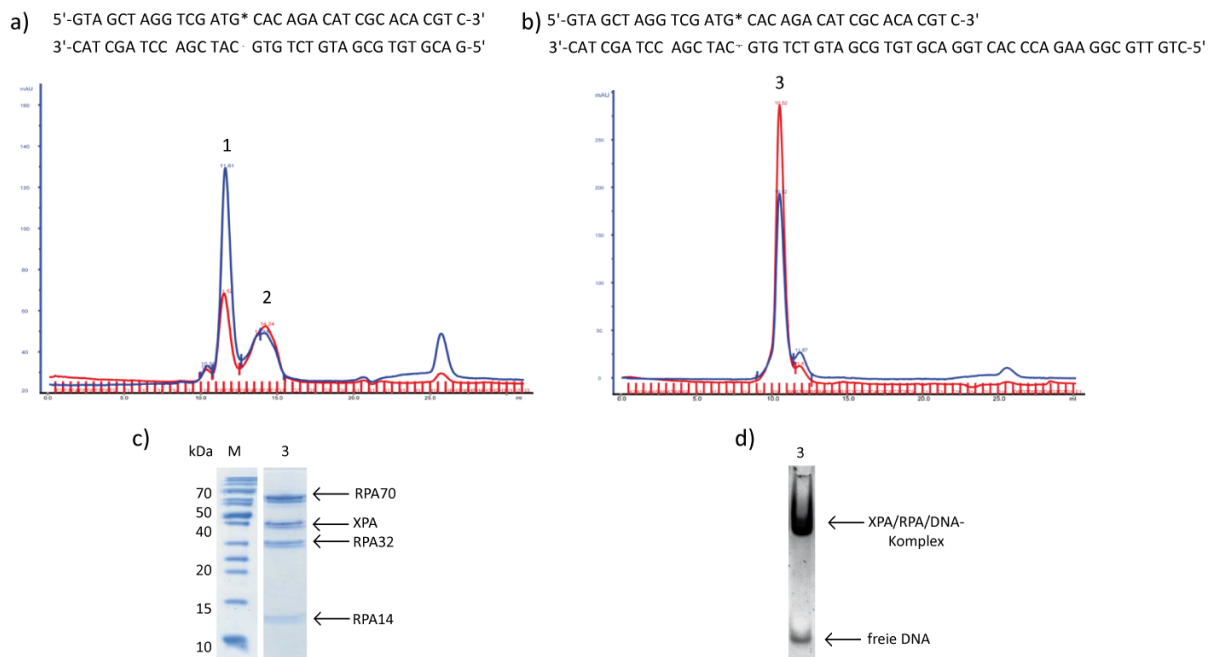


Abbildung 28: Vergleich der Chromatogramme von der XPA/RPA-Probe im Verhältnis von 4:4:1 mit einem AAF-geschädigten 34mer a) ohne Überhang b) mit einem 20nt 5'-Überhang. Sequenz der verwendeten DNA mit G*: AAF(C8)-dG. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL) und Kristallisationspuffer II. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. Die Hauptpeaks beider Chromatogramme haben ein Retentionsvolumen von 11.61 mL (1), 14.24 mL (2), und 10.52 mL (3). **c) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion von 3.** M: Proteinstandard PageRuler™, 3: Eluat bei 10.52 mL. **d) SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 3.** 3: Eluat bei 10.52 mL.

Weiterhin wurden die Protein und DNA enthaltenden Fraktionen der Gelfiltration mit dem Überhang *via* SDS-PAGE und einem nativen Gel analysiert (Abbildung 28c und d). Bei Vergleich beider Chromatogramme ist erkennbar, dass sich diesmal ein neuer Peak mit einem niedrigeren Retentionsvolumen von 10.52 mL und nur ein ganz kleiner Nebenpeak gebildet hat (Abbildung 28b). Auch die Absorption von 260 nm ist bei diesem Peak sehr hoch. Mit dem zuvor verwendeten DNA-Substrat ohne Überhang bildeten sich zwei Peaks mit einem höheren Retentionsvolumen von 11.61 mL und 14.25 mL aus (Abbildung 28a). Dies lässt vermuten, dass sich ein XPA/RPA/DNA-Komplex ausgebildet hat, da sich das Retentionsvolumen zu einem niedrigeren Wert als zuvor verschoben hat. Nach Analyse des neuen Peaks mittels SDS-PAGE (Abbildung 28c) konnten die

Proteine XPA und RPA identifiziert werden. Nach Abschätzung der Bandenintensität enthielt dieser einen 1:1 XPA/RPA-Komplex. Dies bestätigt die Ergebnisse von *Yang et al.*, dass XPA als Monomer an DNA mit ss und ds Bereiche bindet.^[148] Die anschließende Charakterisierung der Fraktion mit einem nativen Gel (Abbildung 28d) zeigte, dass ein DNA-Protein-*Shift* und kaum ungebundene DNA vorlag. Somit bindet der XPA/RPA-Komplex an die DNA. Folglich führte die Optimierung des DNA-Substrats basierend auf den Studien von *Yang et al.* zur Bildung eines XPA/RPA/DNA-Komplexes.^[148] Allerdings wäre der berechnete Wert (421 kDa) für die molekulare Masse des Komplexes (ca. 153 kDa bei einem Komplexverhältnis von 1:1:1) mit der Kalibrationsgeraden aus Abschnitt 4.3.2.3.1. zu hoch. Aufgrund des Inkubationsverhältnis von 4:4:1 (XPA:RPA:DNA) könnten mehrere XPA/RPA-Proteine an die DNA gebunden haben. Daher müssten noch verschiedene Inkubationsverhältnisse sowie eine kleinere DNA getestet werden. Weiterhin könnten aber auch die vielen flexiblen Domänen der Proteine das Laufverhalten des XPA/RPA/DNA-Komplexes ändern.

4.3.4.4. Studien zum XPA/RPA/DNA-Komplex

Nach Isolierung eines XPA/RPA/DNA-Komplexes sollten verschiedene Tests durchgeführt werden, bei denen die Stabilität des Komplexes, die Inkubationszeit, die Komplexverhältnisse, sowie die Inkubationszeiten der einzelnen Komponenten variiert werden. Für die Stabilitätsstudien wurden die Fraktionen des isolierten Komplexes einkonzentriert und erneut auf die Gelfiltrationssäule aufgetragen. Abbildung 29 stellt den Stabilitätstests dar.

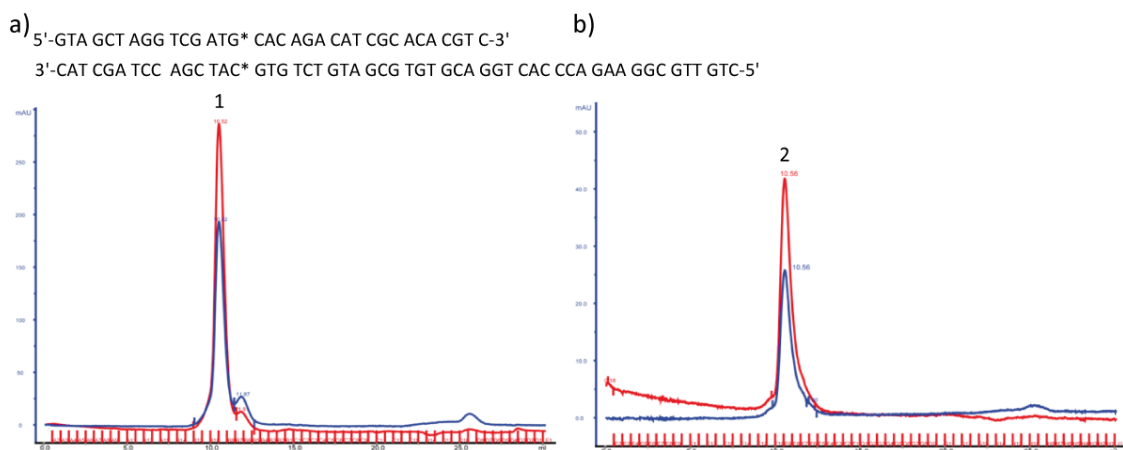


Abbildung 29: Stabilitätstests des XPA/RPA/DNA-Komplexes (4:4:1) mit einem AAF-geschädigten 34mer mit einem 20nt 5'-Überhang. a) Isolierung des Komplexes. b) Erneute Gelfiltration des isolierten Komplexes von Peak 1 aus a. Sequenz der verwendeten DNA mit G*: AAF(C8)-dG. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL) und Kristallisationspuffer II. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. Die Hauptpeaks beider Chromatogramme haben ein Retentionsvolumen von 10,52 mL (1) und 10,56 mL (2).

Nach Isolierung des Komplexes (29a) konnte nach erneuter Gelfiltration des isolierten Komplexes wieder ein Peak bei einem gleichen Retentionsvolumen (10.56 mL, Abbildung 29b) erhalten werden. SDS-PAGE und native Gele bestätigen weiterhin die Komplexstabilität. Zudem wurde noch getestet wie lange der Komplex stabil ist. Dafür wurde der Komplex isoliert und bei 4 °C oder bei -80 °C über Nacht gelagert. Anschließend wurden diese erneut auf die Gelfiltrationssäule aufgetragen. Es konnte ein Peak bei demselben Retentionsvolumen, wie zuvor schon in Abbildung 29b, erhalten werden. Diese Ergebnisse zeigen, dass der isolierte XPA/RPA/DNA-Komplex sehr stabil ist und sogar nach Lagerung bei -80 °C beständig bleibt.

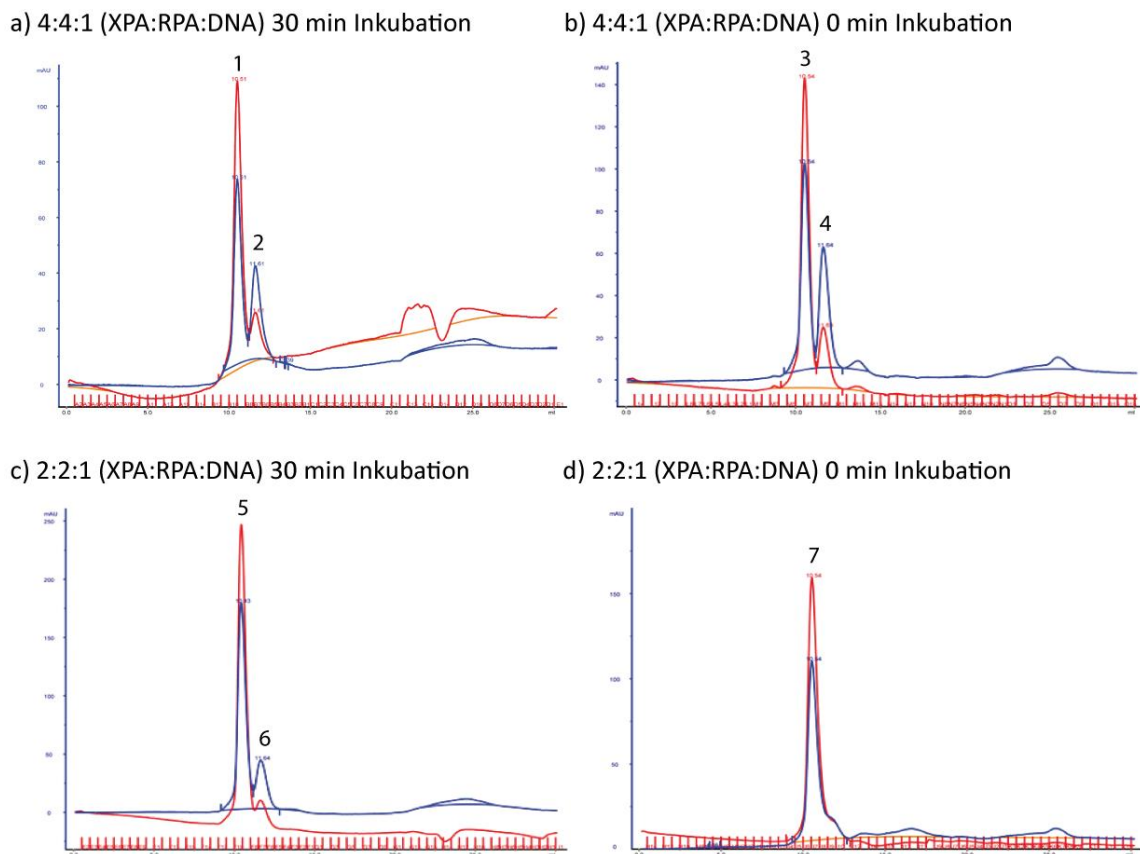


Abbildung 30: Vergleich der verschiedenen Komplexverhältnisse und Inkubationszeiten mit einem AAF-geschädigten 34mer mit einem 20nt 5'-Überhang. a) Komplexverhältnis von 4:4:1 und 30 min Inkubationszeit. b) Komplexverhältnis von 4:4:1 und 0 min Inkubationszeit. c) Komplexverhältnis von 2:2:1 und 30 min Inkubationszeit. d) Komplexverhältnis von 2:2:1 und 0 min Inkubationszeit. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL) und Kristallisationspuffer II. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm.

Ferner sollten noch verschiedene Komplexverhältnisse sowie Inkubationszeiten getestet werden, um das beste Verhältnis und die optimale Inkubationszeit zu finden. Dafür wurde der Komplex in einem Verhältnis von 4:4:1 (XPA:RPA:DNA) und 2:2:1 für 30 min oder 0 min bei 4 °C inkubiert. Abbildung 30 zeigt die Ergebnisse dieser Studie. Auffällig war, dass nach mehrmaliger Wiederholung der Komplexbildung bei einem Verhältnis von 4:4:1 (XPA:RPA:DNA) teilweise ein kleiner Nebenpeak entsteht (Abbildung 30a, Peak 2), welcher einen XPA/RPA-Komplex enthielt. Zudem ist kein Unterschied bei

einer Komplexinkubation von 30 min oder 0 min (Abbildung 30a und b) zu erkennen. Darum könnte eine Erhöhung der DNA-Konzentration den Nebenpeak reduzieren. Abbildung 30c zeigt die Komplexbildung bei einem Verhältnis von 2:2:1 bei einer Inkubation von 30 min. Dabei konnte die gleiche Komplexbildung wie bei einem Verhältnis von 4:4:1 mit gleichen Retentionsvolumen erhalten werden. Zudem ist der Nebenpeak mit einem XPA/RPA-Komplex etwas kleiner (Peak 6). Erst die Reduzierung der Inkubationszeit auf 0 min ermöglicht die Entfernung des Nebenpeaks und somit eine vollständige Komplexbildung (Abbildung 30d). Damit konnte die Inkubationszeit reduziert werden, da sich der Komplex sofort ausbilden kann. Für alle weiteren Untersuchungen wurde das Komplexverhältnis von 2:2:1 weiterverwendet.

Verschiedene Inkubationszeiten der einzelnen Komponenten sollten im nächsten Schritt evaluiert werden, um zu testen, ob die Komplexbildung erhalten bleibt. Abbildung 31 zeigt die verschiedenen Zugabezeiten und die anschließende Analyse des Komplexes mittels Gelfiltration.

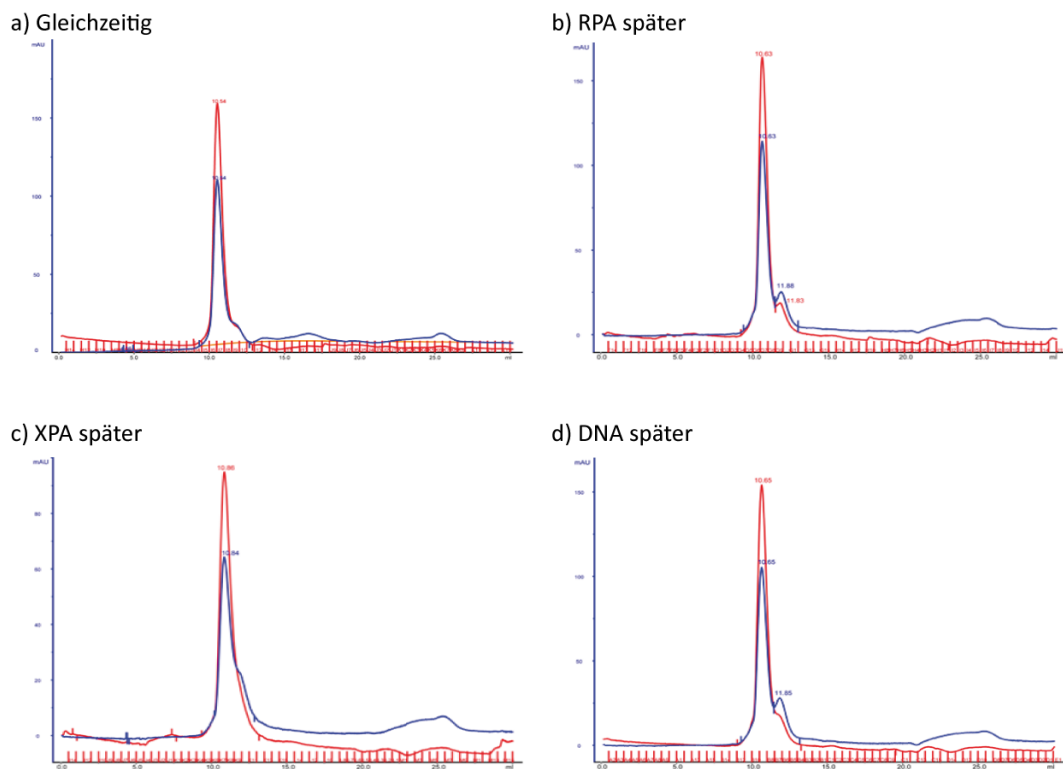


Abbildung 31: Vergleich der verschiedenen Zugabezeiten der Proteine und einem AAF-geschädigten 34mer mit einem 20nt 5'-Überhang zur Komplexbildung (Verhältnis 2:2:1 XPA:RPA:DNA). a) Gleichzeitig und 0 min Inkubation. b) Inkubation von XPA und DNA für 10 min; dann Zugabe von RPA. c) Inkubation von RPA und DNA für 10 min; dann Zugabe von XPA. d) Inkubation von XPA und RPA für 10 min; dann Zugabe von DNA. Verwendung der Gelfiltrationssäule SuperdexTM 200 Increase 10/300 GL (24 mL) und Kristallisationspuffer II. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm.

Für die verschiedenen Inkubationszeiten konnte stets ein XPA/RPA/DNA-Komplex erhalten werden. Teilweise bildete sich ein kleiner Nebenpeak mit XPA/RPA aus. Dies bedeutet, dass es keine Rolle spielt in welcher Reihenfolge die Proteine und DNA zur Komplexbildung hinzugefügt werden. Zum

anderen könnte das auch bedeuten, dass sich vorweg wahrscheinlich kein XPA/RPA-Komplex ausbildet, da die Reihenfolge der Proteine zur Komplexbildung keinen Unterschied macht. Diese Hypothese lässt sich mit den EMSA-Studien von *Yang et al.* unterstützen.^[148] Dabei nehmen diese an, dass die Proteine erst bei DNA Bindung miteinander interagieren.^[148]

4.3.4.5. Screening von DNA-Sequenzen für die Komplexbildung

4.3.4.5.1. Die verschiedenen DNA-Substrate

Nach Optimierung des Gelfiltrationsassays sollten verschiedene ds DNA-Sequenzen getestet werden, wie eine kürzere DNA, ein kürzerer Überhang, ohne DNA-Schaden oder eine *Bubble*-DNA, um zu sehen, wo sich noch ein XPA/RPA/DNA-Komplex ausbildet. Theoretisch könnte das bisher verwendete 34mer oder der 20nt Überhang zu lang sein, wodurch nicht alles von XPA/RPA gebunden sein könnte. Dadurch hätte die DNA flexible freie Bereiche, was die Kristallisation erschweren würde. Daher war es wichtig einen kürzeren Überhang als auch eine kürzere DNA zu testen. Auch wäre es interessant zu wissen, ob sich auch ein Komplex mit ungeschädigter DNA ausbildet oder ob der Komplex Schaden-spezifisch bindet. Ein wichtiger Aspekt ist die *Bubble*-DNA als Substrat, da diese dem natürlichem Intermediat, dem Prä-Inzisionskomplex, während der NER am nächsten kommt. Tabelle 20 listet die über 25 getesteten DNA-Substrate für die Komplexbildung auf.

Tabelle 20: Auflistung der verwendeten DNA-Substrate zur Komplexbildung. Substrate, die zu einer Komplexbildung führten, sind blau hervorgehoben.

ds DNA-Substrate	Zusatz	DNA-Schaden
15mer	-	AAF(C8)-dG
	-	Ungeschädigt
	ss 20nt 5'-Überhang	AAF(C8)-dG
25mer	-	AAF(C8)-dG
	-	Ungeschädigt
	ss 20nt 5'-Überhang	AAF(C8)-dG
	ss 20nt 5'-Überhang	Ungeschädigt
	ss 15nt 5'-Überhang	AAF(C8)-dG
	ss 15nt 5'-Überhang	Ungeschädigt
	ss 10nt 5'-Überhang	AAF(C8)-dG
	2Y-DNA, jeweils an beiden Enden 7 ungepaarte Fehlstellen	AAF(C8)-dG
33mer	13mer Bubble durch 13 ungepaarte Fehlstellen	AAF(C8)-dG
	13mer Bubble durch 13 ungepaarte Fehlstellen	Ungeschädigt
34mer	-	AAF(C8)-dG
	-	Ungeschädigt
	ss 20nt 5'-Überhang	AAF(C8)-dG
	ss 20nt 5'-Überhang	Ungeschädigt
	ss 15nt 5'-Überhang	AAF(C8)-dG
	ss 10nt 5'-Überhang	AAF(C8)-dG
	ss 1nt 5'-Überhang oder 1nt 3'-Überhang	AAF(C8)-dG
	An beiden Enden ein ss 20nt 5'-Überhang	AAF(C8)-dG
	13mer Bubble durch 13 ungepaarte Fehlstellen	AAF(C8)-dG
	13mer Bubble durch 13 ungepaarte Fehlstellen	Ungeschädigt
39mer	19mer Bubble durch 19 ungepaarte Fehlstellen	AAF(C8)-dG
	19mer Bubble durch 19 ungepaarte Fehlstellen	Ungeschädigt
45mer_1	25mer Bubble durch 25 ungepaarte Fehlstellen	Ungeschädigt

Die besten DNA-Substrate, die eine ausreichende Komplexbildung aufwiesen, wurden für weitere strukturelle Studien verwendet und sind in der Tabelle 20 blau hervorgehoben. Im Folgenden

werden nur diese DNA-Substrate beschrieben und diskutiert. Alle weiteren Chromatogramme, SDS-PAGE, native Gele, Retentionsvolumen der einzelnen DNA, sowie die Chromatogramme einzelner Proteine mit den DNA-Substraten, befinden sich im Anhang 6.5. Ferner befinden sich EMSA-Studien im Anhang 6.6.

4.3.4.5.2. DNA Substrat: 34mer mit einem 20nt 5'-Überhang

In den Abschnitten zuvor (4.3.4.3.3. und 4.3.4.4.) konnte das AAF-geschädigte 34mer mit einem 20nt 5'-Überhang charakterisiert werden, welches zu einer guten Komplexbildung führte. Auch das ungeschädigte 34mer mit einem 20nt 5'-Überhang führte zu einer ähnlichen Komplexbildung. Abbildung 32 zeigt den Vergleich beider Substrate. Aufgrund der gleichen Retentionsvolumina handelt es sich vermutlich um den gleichen Komplex (Abbildung 32a und b). Dies wurde auch durch die SDS-PAGE, sowie durch das native Gel bestätigt (Abbildung 32c und d). Es bildet sich ein XPA/RPA/DNA-Komplex mit einem 34mer und einem 20nt 5'-Überhang sowohl ohne Schaden als auch mit einem AAF-Schaden aus. Daraus kann geschlossen werden, dass XPA die DNA an der ds/ss-Verzweigung erkennt bzw. bindet und nicht den Schaden selbst erkennt. Bereits *Yang et al.* zeigten, dass XPA stärker an verzweigte- (Y-förmige DNA) oder *Bubble*-DNA als an geschädigter DNA binden kann.^[148]

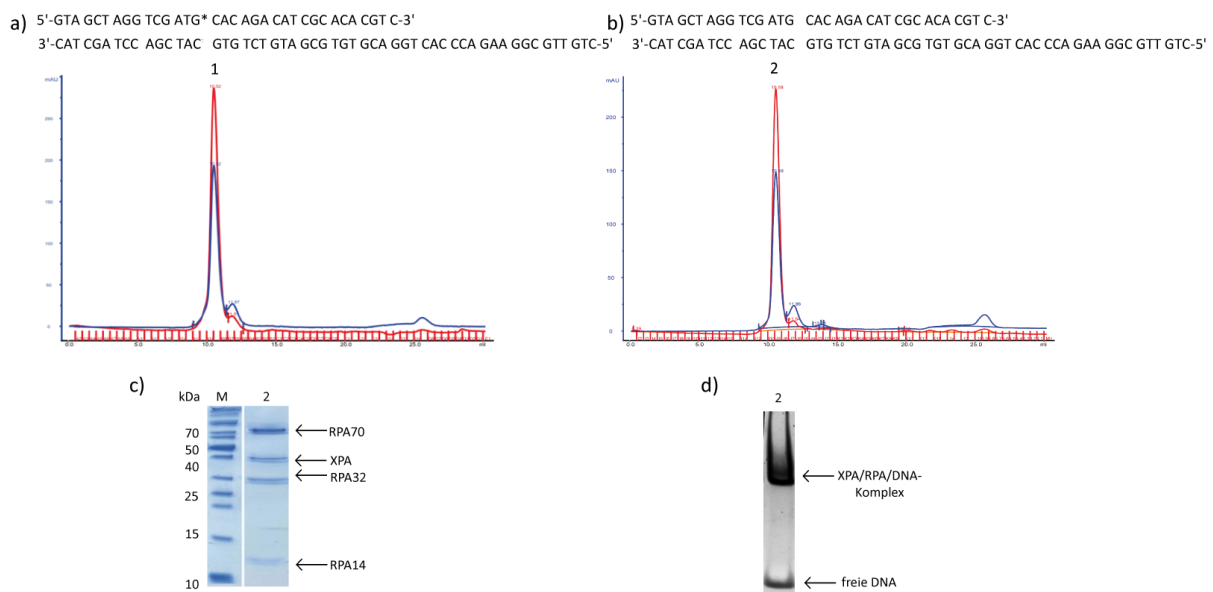


Abbildung 32: Vergleich der XPA/RPA-Probe im Verhältnis von 2:2:1 mit einem 34mer mit einem ss 20nt 5'-Überhang a) mit G*: AAF(C8)-dG Schaden b) ungeschädigt. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL), Kristallisationspuffer II und 0 min Inkubationszeit. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. Die Hauptpeaks beider Chromatogramme haben ein Retentionsvolumen von 10,52 mL (1) und 10,56 mL (2). **c) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion von 2.** M: Proteinstandard PageRuler™, 2: Eluat bei 10,56 mL. **d) SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 2.** 2: Eluat bei 10,56 mL.

Außerdem wurden noch Überhänge verschiedener Größe getestet. 5'-Überhänge von 1nt und 10nt führten nicht zur Komplexbildung. Abbildung 33 stellt die Komplexbildung mit einem AAF-geschädigten 34mer und einem 15nt 5'-Überhang dar. Dabei bildet sich zwar ein vollständiger Komplex, da nur ein Peak entsteht (Abbildung 33a). Die Komplexbildung konnte durch SDS-PAGE und natives Gel verifiziert werden (Abbildung 33b und c). Dennoch befindet sich im nativen Gel hauptsächlich freie DNA und wenig Protein-DNA-Shift. Dies könnte bedeuten, dass sich zwar ein Komplex bildet, aber aufgrund seiner Instabilität wieder zerfällt.

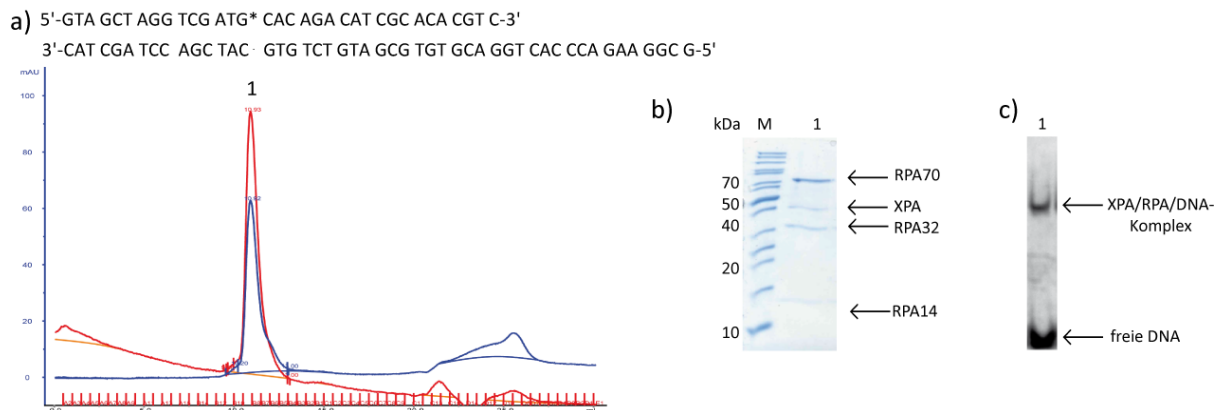


Abbildung 33: a) Chromatogramm des Gelfiltrationsassays mit XPA zusammen mit RPA und ein AAF-geschädigtes 34mer mit einem 15nt 5'-Überhang im Verhältnis von 2:2:1. G*: AAF(C8)-dG Schaden. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL). x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. Der Hauptpeak hat ein Retentionsvolumen von 10.93 mL (1). b) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion von 1. M: Proteinstandard PageRuler™, 1: Eluat bei 10.93 mL. c) SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 1. 1: Eluat bei 10.93 mL.

RPA kann längere DNA stabiler binden (20-30 nt Bindungsmodus) als kürzere DNA (8-10nt Bindungsmodus).^[57] Vermutlich bindet RPA bei der Komplexbildung an den ss Überhang, wobei ein 20nt Überhang eine optimale Länge zur Komplexbildung darstellt. Sobald der Überhang kleiner wird, bindet RPA schwächer, bis gar nicht. Dadurch kommt es zu einer schwächeren oder gar keinen Komplexbildung.

4.3.4.5.3. DNA Substrat: 25mer mit einem 20nt 5'-Überhang

Als nächstes sollte eine kleinere DNA, ein 25mer, mit einem 20nt 5'-Überhang für die Komplexbildung getestet werden. Bei Komplexbildung könnten die dann zuvor im 34mer freien Bereiche der DNA reduziert werden, was die Kristallisation erleichtern würde. Abbildung 34 zeigt den Vergleich vom 25mer mit und ohne AAF-Schaden. Dabei kann aus beiden Chromatogrammen entnommen werden, dass sich für beide Substrate ein Komplex ausbildete (Abbildung 34a und b). Lediglich ein kleiner Nebenpeak entstand. Die Komplexbildung konnte auch durch die SDS-PAGE,

sowie durch das native Gel, bestätigt werden (Abbildung 34c, d, e und f). Zudem liegen XPA und RPA in einem 1:1 Verhältnis nach Abschätzung der Bandenintensität im SDS-PAGE (Abbildung 34c und e) vor. Die Resultate stimmen mit den zuvor erhaltenen Ergebnisse des 34mers mit einem 20nt 5'-Überhang ein. Der XPA/RPA/DNA-Komplex bildet sich sowohl ohne Schaden als auch mit einem AAF-Schaden aus, da XPA vermutlich den ds/ss-Übergang und RPA an den ss Überhang bindet.

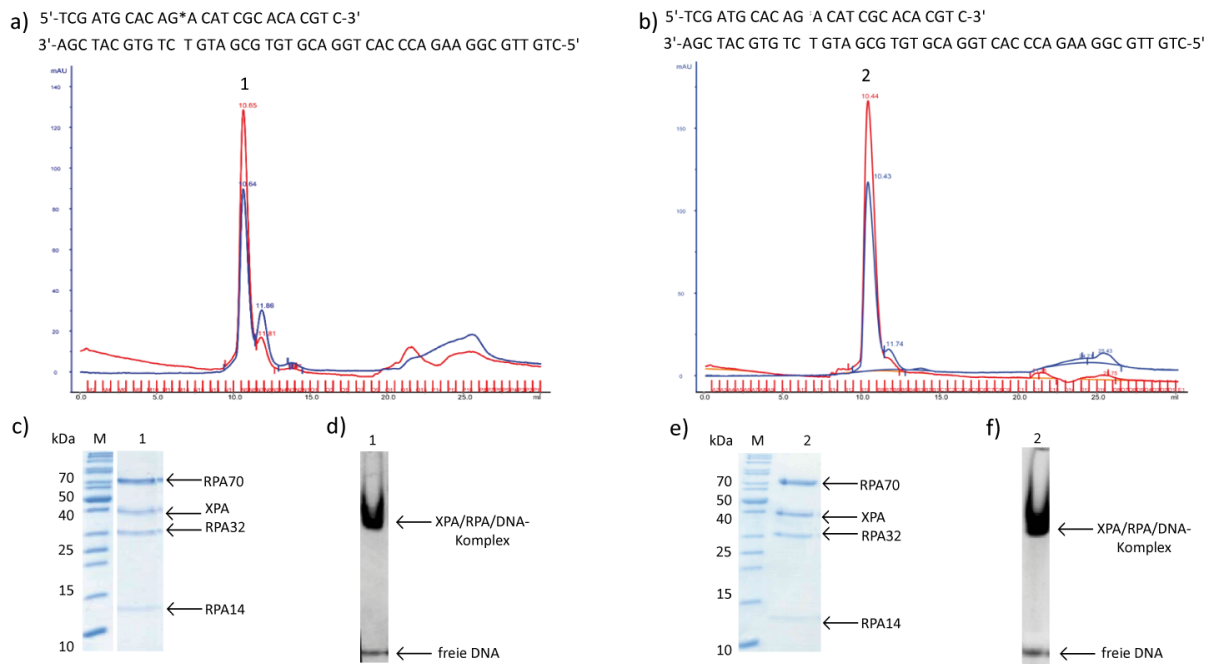


Abbildung 34: Vergleich der XPA/RPA-Probe im Verhältnis von 2:2:1 mit einem 25mer und einem ss 20nt 5'-Überhang a) mit G*: AAF(C8)-dG Schaden b) ungeschädigt. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL), Kristallisationspuffer II und 0 min Inkubationszeit. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. Die Hauptpeaks beider Chromatogramme haben ein Retentionsvolumen von 10.65 mL (1) und 10.54 mL (2). **c) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion von 1.** M: Proteinstandard PageRuler™, 1: Eluat bei 10.65 mL. **d) SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 1.** 1: Eluat bei 10.65 mL. **e) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion von 2.** M: Proteinstandard PageRuler™, 2: Eluat bei 10.54 mL. **f) SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 2.** 2: Eluat bei 10.54 mL.

Ebenso ist der Komplex mit dem 25mer und 20nt 5'-Überhang, wie das 34mer, bei 4 °C über Nacht oder bei -80 °C stabil. Weiterhin wurde auch ein 15mer mit einem 20nt 5'-Überhang als DNA-Substrat überprüft. Dabei bildete sich kaum ein Komplex aus. Dieses Substrat ist vermutlich für die Ausbildung des Komplexes zu klein, sodass die Proteine nicht gut binden können.

Außerdem wurden, wie beim 34mer, verschiedene Überhänge (15nt und 10nt) zur Komplexbildung getestet, wobei die erhaltenen Ergebnisse gleich sind. Abbildung 35 stellt die Komplexbildung mit einem AAF-geschädigten 25mer und einem 15nt 5'-Überhang dar. Dabei bildet sich zwar ein vollständiger Komplex, da nur ein Peak entsteht (Abbildung 35a), aber wie beim 34mer ist dieser nicht stabil. Die Komplexbildung konnte durch SDS-PAGE und natives Gel verifiziert werden (Abbildung 35b und c). Da sich im nativen Gel hauptsächlich freie DNA und kaum ein Protein-DNA-Shift befand, könnte der entstandene Komplex zerfallen sein. Mit dem 10nt 5'-Überhang konnte

keine Komplexbildung erzielt werden. Folglich konnte ein etwas kleineres Substrat (25mer) für die Komplexbildung gefunden werden. Hinzu konnte bestätigt werden, dass der 20nt 5'-Überhang die optimale Länge zur Komplexbildung aufweist.

a) 5'-TCG ATG CAC AG*A CAT CGC ACA CGT C-3'

3'-AGC TAC GTG TC T GTA GCG TGT GCA GGT CAC CCA GAA GGC G-5'

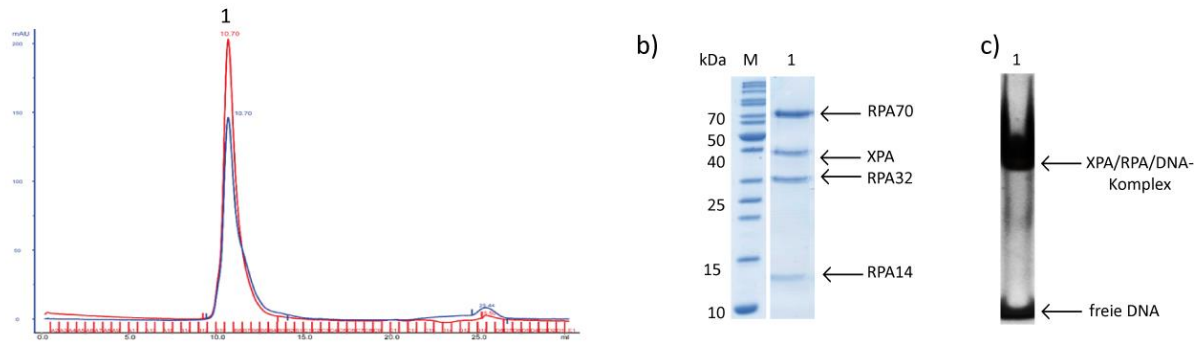


Abbildung 35: a) Chromatogramm des Gelfiltrationsassays mit XPA zusammen mit RPA und ein AAF-geschädigtes 25mer mit einem 15nt 5'-Überhang im Verhältnis von 2:2:1. G*: AAF(C8)-dG Schaden. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL). x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. Der Hauptpeak hat ein Retentionsvolumen von 10.70 mL (1). b) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion von 1. M: Proteinstandard PageRuler™, 1: Eluat bei 10.70 mL. c) SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 1. 1: Eluat bei 10.70 mL.

4.3.4.5.4. DNA-Substrat: AAF-geschädigte *Bubble*-DNA

Zuletzt wurde eine *Bubble*-DNA getestet, da diese dem natürlichen Intermediat, dem Prä-Inzisionskomplex, während der NER, näher kommt. Dabei sollte eine 33mer mit einer 13nt *Bubble*, welche durch 13 Basenpaar-*Mismatches* entsteht, verwendet werden. Abbildung 36 zeigt den Vergleich vom 33mer mit und ohne AAF-Schaden. Dabei kann aus beiden Chromatogrammen entnommen werden, dass sich unterschiedliche Peaks bildeten (Abbildung 36a und b). Dabei entstand ein Hauptpeak mit einem Retentionsvolumen von 9.75 mL (Abbildung 36a, Peak 1) für das geschädigte Substrat und lediglich ein kleiner Nebenpeak. Für das ungeschädigte Substrat bildeten sich mehrere Peaks, wobei der intensivste ein Retentionsvolumen von 10.00 mL hatte (Abbildung 36b, Peak 3) und somit nicht den gleichen Wert wie das geschädigte Substrat aufwies. Daher wurden die Peaks zunächst durch SDS-PAGE und durch ein natives Gel analysiert. Dabei konnte die Komplexbildung für die geschädigte *Bubble* bestätigt werden (36c und d, Peak 1). Der Nebenpeak enthielt einen XPA/RPA-Komplex mit sehr wenig DNA (36c und d, Peak 2). Im Gegensatz dazu konnte keine Komplexbildung für die ungeschädigte *Bubble* erhalten werden. Interessanterweise zeigt der Peak 3 und 4 nur einen RPA/DNA-Komplex (36e und f, Peak 3 und 4). Wahrscheinlich haben mehr RPA-Proteine im Peak 3 an die DNA gebunden als an Peak 4. Dies würde die unterschiedlichen

Retentionsvolumina erklären. Daneben bildet sich noch ein instabiler XPA/DNA-Komplex aus (36e und f, Peak 5), sowie ein DNA-Komplex (36e und f, Peak 6) aus. Diese Ergebnisse sind von sehr hohem Interesse für die aktuelle Forschung, da sich der Komplex nur mit einem DNA-Schaden in der *Bubble* ausbilden kann.

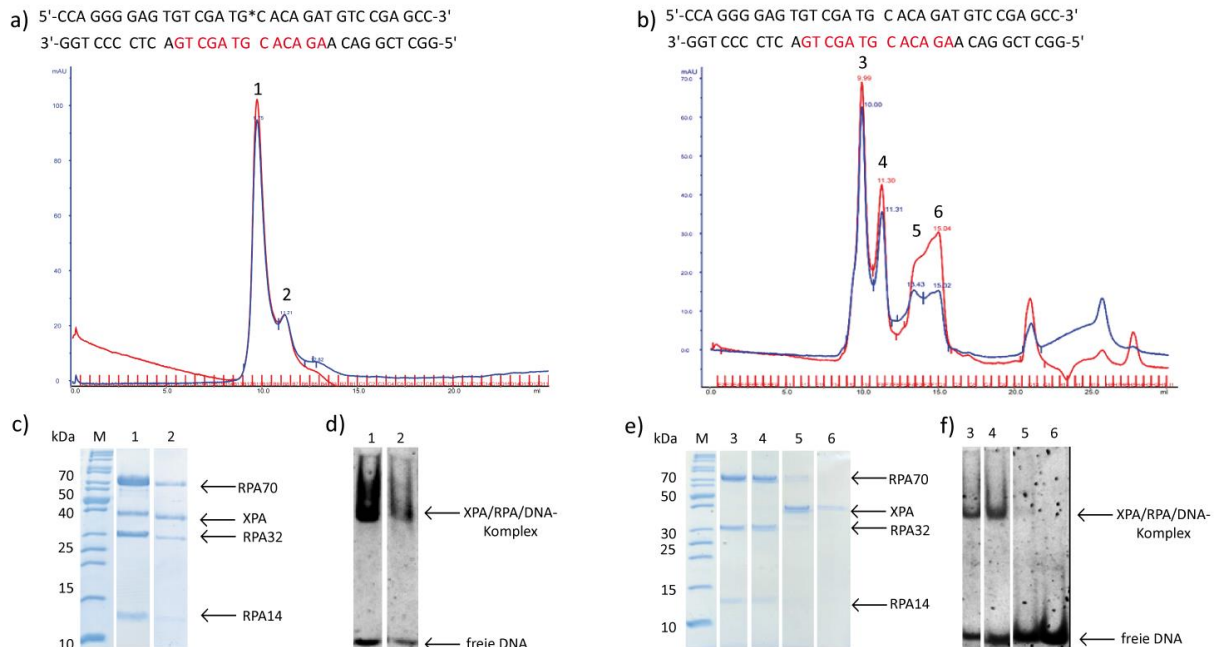


Abbildung 36: Vergleich der XPA/RPA-Probe im Verhältnis von 2:2:1 mit einer 13nt *Bubble* im 33mer, welche durch 13 ungepaarte Fehlstellen (in rot markiert) entsteht. a) mit G*: AAF(C8)-dG Schaden b) ungeschädigt. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL), Kristallisationspuffer II und 0 min Inkubationszeit. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. Die Hauptpeaks beider Chromatogramme haben ein Retentionsvolumen von 9.75 mL (1), 11.21 (2), 10.00 (3), 11.31 (4), 13.43 (5) und 15.02 mL (6). **c) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion von 1 und 2.** M: Proteinstandard PageRuler™, 1: Eluat bei 9.75 mL, 2: Eluat bei 11.21 mL. **d) SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 1 und 2.** **e) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion von 3, 4, 5 und 6.** M: Proteinstandard PageRuler™, 3: Eluat bei 10.00 mL, 4: Eluat bei 11.31 mL, 5: Eluat bei 13.43 mL, 6: Eluat bei 15.02 mL. **f) SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 3, 4, 5 und 6.**

Bei Betrachtung des Komplexes mit einer AAF-geschädigten Bubble-DNA im SDS-PAGE (Abbildung 36c) fällt auf, dass XPA und RPA in einem 1:2 Verhältnis nach Abschätzung der Bandenintensität vorliegen. Im Vergleich zum 34mer und 25mer (2:1 XPA:RPA) ist das Verhältnis genau umgekehrt. Studien von *Chen et al.* zeigen, dass mehrere RPA Proteine notwendig sind, um einen stabilen Komplex mit einer Bubble-DNA auszubilden.^[57] Diese Studien könnten das XPA:RPA Verhältnis von 1:2 erklären. Das mit dem Standard berechnete Molekulargewichtes (647 kDa) des Komplexes gibt keinen Hinweis auf die Zusammensetzung des Komplexes. Vermutlich verändert sich das Laufverhalten des Komplexes durch die ungepaarten Basen in der *Bubble*, wodurch der Komplex eher eluieren kann.

Da die *Bubble*-DNA dem natürlichen Substrat in der NER sehr ähnlich ist, wurden verschiedene Tests durchgeführt werden, bei denen die Stabilität des Komplexes, die Komplexverhältnisse, sowie die

Zugabereihenfolge der einzelnen Komponenten variiert wurden. Abbildung 37 zeigt die Ergebnisse der verschiedenen Komplexverhältnisse. Die Komplexbildung bei einem Verhältnis von 4:4:1 (37a, Peak 1), zeigte noch einen kleinen RPA/XPA-Komplex Nebenpeak (Peak 2). Dieser konnte durch das Komplexverhältnis 2:2:1 reduziert werden (37b, Peak 4). Durch Erhöhung der DNA (37c), konnte der Nebenpeak nicht entfernt werden, sondern ein zusätzlicher DNA Peak entstand (Peak 7). Damit zeigt das Verhältnis 2:2:1 die beste Komplexbildung.

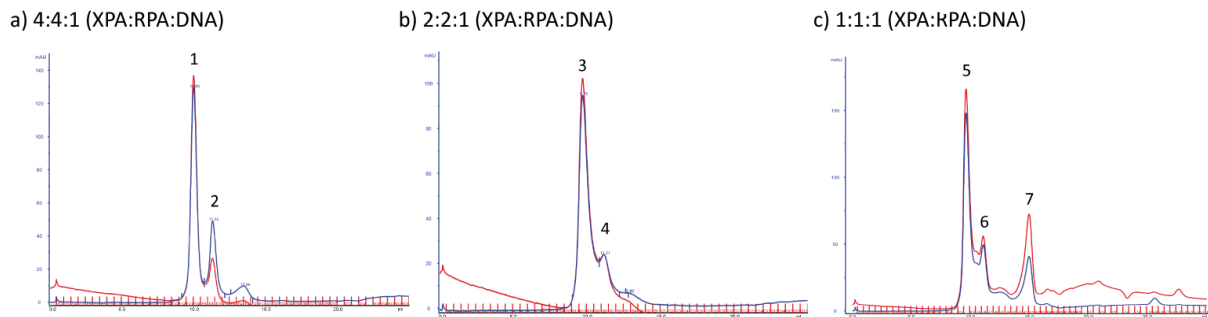


Abbildung 37: Vergleich der verschiedenen XPA/RPA-Komplexverhältnisse mit einer AAF-geschädigten 13nt *Bubble* in einem 33mer. a) Komplexverhältnis von 4:4:1 (XPA:RPA:DNA). b) Komplexverhältnis von 2:2:1 (XPA:RPA:DNA). c) Komplexverhältnis von 1:1:1 (XPA:RPA:DNA). Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL) und Kristallisationspuffer II. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm.

Für die Stabilitätsstudien wurden die Fraktionen des isolierten Komplexes einkonzentriert und erneut auf die Gelfiltrationssäule aufgetragen. Abbildung 38 stellt den Stabilitätstest dar.

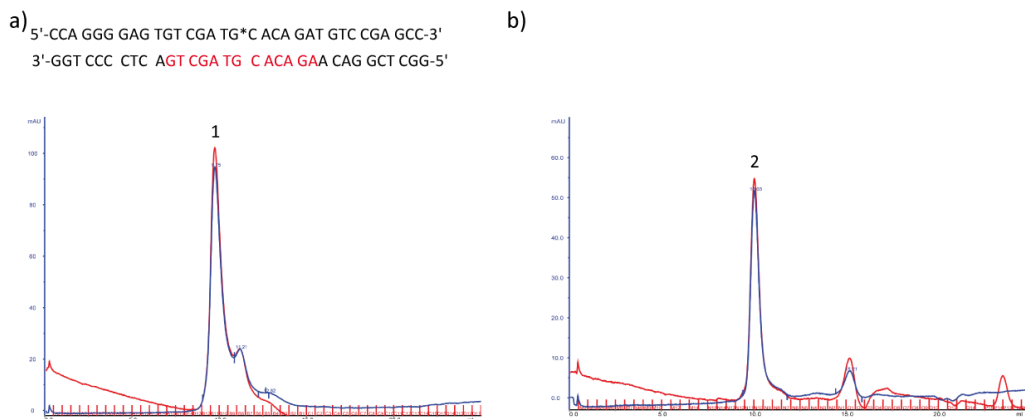


Abbildung 38: Stabilitätstests des XPA/RPA/DNA-Komplexes (2:2:1) mit einer AAF-geschädigten 13nt *Bubble* im 33mer. a) Isolierung des Komplexes. b) Erneute Gelfiltration des isolierten Komplexes von Peak 1 aus a. Sequenz der verwendeten DNA mit G*: AAF(C8)-dG. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL) und Kristallisationspuffer II. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. Die Hauptpeaks beider Chromatogramme haben ein Retentionsvolumen von 9.75 mL (1) und 10.03 mL (2).

Nach Isolierung des Komplexes (38a) konnte nach erneuter Gelfiltration der gleiche Komplex erhalten werden (38b). SDS-PAGE und native Gele bestätigen weiterhin die Komplexbildung und

befinden sich im Anhang 6.5. Zudem ist der isolierte Komplex bei 4 °C über Nacht oder bei Lagerung bei -80 °C beständig.

Zum Schluss sollten noch die verschiedene Inkubationszeiten der einzelnen Komponenten evaluiert werden, um zu testen, ob die Komplexbildung erhalten bleibt. Abbildung 39 zeigt die verschiedenen Zugabezeiten und die anschließende Analyse des Komplexes mittels Gelfiltration.

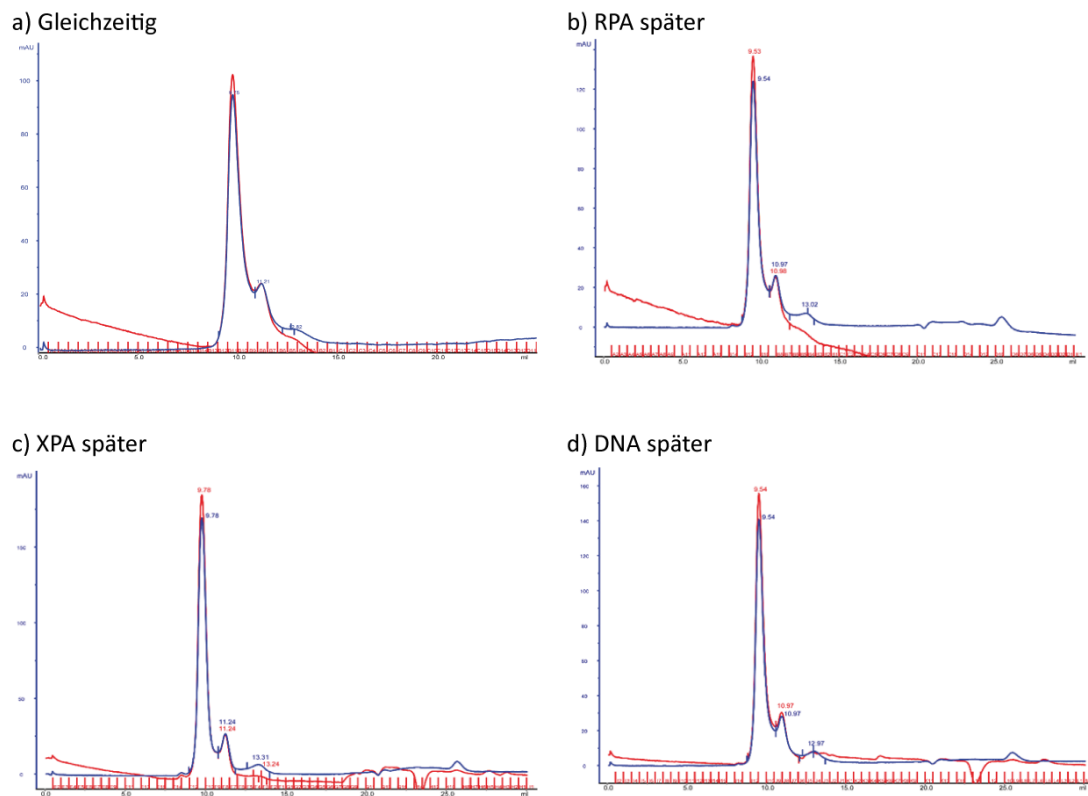


Abbildung 39: Vergleich der verschiedenen Zugabezeiten der Proteine und einer AAF-geschädigten 13nt *Bubble* im 33mer zur Komplexbildung (Verhältnis 2:2:1 XPA:RPA:DNA). a) Gleichzeitig und 0 min Inkubation. b) Inkubation von XPA und DNA für 10 min; dann Zugabe von RPA. c) Inkubation von RPA und DNA für 10 min; dann Zugabe von XPA. d) Inkubation von XPA und RPA für 10 min; dann Zugabe von DNA. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL) und Kristallisationspuffer II. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm.

Dabei konnte für alle Zugabezeiten ein XPA/RPA/DNA-Komplex erhalten werden, sodass die unterschiedliche Reihenfolge der Proteine die Komplexbildung nicht beeinträchtigt.

Außerdem wurden noch verschiedene *Bubble*-DNAs getestet. Zum einen ein 34mer mit einer 13nt *Bubble*, wobei der AAF-Schaden nicht ganz symmetrisch in der Mitte lokalisiert war. Dabei sollte vor allem überprüft werden, ob der Komplex auch bei einer anderen Sequenz entsteht. Zum anderen ein 39mer mit einer 16nt *Bubble*, um eine Komplexbildung mit einer größeren *Bubble* zu erhalten. Für beide Substrate konnte wieder ein Komplex nur dem DNA-Schaden, wie beim 33mer, erhalten werden.

Folglich konnte der Gelfiltrationsassay die stabile XPA/RPA/DNA-Komplexbildung mit verschiedenen DNA-Substraten nachweisen. Allerdings lieferte dieser keine exakte Aussage über die Dynamik des Komplexes. Denn z.B. könnte XPA zuerst an die DNA binden und anschließend RPA rekrutieren, wodurch RPA die DNA binden kann. Interessant wäre dann auch, ob RPA nach DNA-Bindung noch mit allen XPA-Interaktionsstellen mit XPA interagiert oder ob es zu einer strukturellen Umlagerung von RPA kommt. Die Dynamik des Komplexes könnte in Zukunft durch die Kombination von *Crosslinker*- und Strukturstudien untersucht werden.

4.3.5. Crosslinking Studien

Crosslinking Studien ermöglichen unter anderem die Identifizierung von Protein-Protein Interaktionen. Dabei wurden Studien zur Identifizierung von XPA/RPA-Interaktionen zur Strukturaufklärung von *Michael Stadlmeier* durchgeführt. Hierbei wurde XPA zusammen mit RPA und doppelsträngiger AAF-geschädigter oder ungeschädigter DNA (34mer und 15mer) mit einem 20nt 5'-Überhang (ODN: 34mer_1/2 und 34mer_4, 15mer_1/2 und 15mer_4) bei 0 °C für 5 min inkubiert. Die DNA konnte zuvor durch den Gelfiltrationsassay als ein gutes Substrat zur Komplexbildung identifiziert werden. Anschließend wurde dieser Komplex mit dem von *Michael Stadlmeier* entwickelten HSAK+ *Crosslinker*, welcher an zwei Lysine kovalent binden kann, bei 0 °C für 1.5 h inkubiert. Nach enzymatischem Trypsin Verdau der Proben wurden diese mittels HPLC-MS/MS (Ultimate 3000 RSLCnano, gekoppelt an Q Exactive HF) analysiert. Die erhaltenen Daten wurden mit dem Programm MeroX ausgewertet.^[172] Abbildung 40 zeigt den HSAK+ *Crosslinker* und stellt schematisch die erhaltenen Intra-und Inter-*Crosslinks* durch den Webservice xVis dar.^[173] Für diese ersten Ergebnisse mussten mehrere Messungen geschädigter und ungeschädigter DNA-Stränge zusammengefasst werden, da einzelne Messungen eine unzureichende Anzahl an *Crosslinks* zeigten.

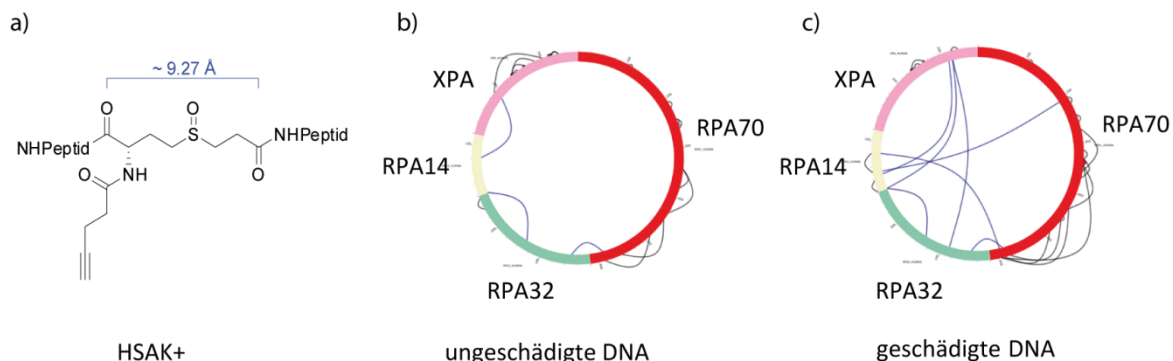


Abbildung 40: a) Darstellung des HSAK+ *Crosslinker*. Schematische Darstellung von Intra- und Inter-*Crosslinks* von b) XPA/RPA-Komplex mit ungeschädigter DNA (15mer und 34mer mit 20nt 5'-Überhang), 8 Intra- und 3 Inter-*Crosslinks*; und c) XPA/RPA-Komplex mit geschädigter DNA (15mer und 34mer mit 20nt 5'-Überhang); 16 Intra- und 7 Inter-*Crosslinks*. Der HSAK+ *Crosslinker* von Michael Stadlmeier wurde mit einem Überschuss von 500 äq. mit den Protein/DNA-Komplex für 1.5 h inkubiert. Die Visualisierung der *Crosslinks* erfolgte durch Michael Stadlmeier mit dem xVis Webserver von Franz Herzog.^[173]

Die Ergebnisse zeigen, dass der XPA/RPA-Komplex mit geschädigter DNA mehr *Crosslinks* als mit ungeschädigter DNA aufweist. Dies bedeutet vor allem, dass der XPA/RPA-Komplex mit geschädigter DNA wahrscheinlich eine andere Konformation einnimmt als bei ungeschädigter DNA. Um weitere Aussagen treffen zu können, müssen die Experimente, insbesondere mit einer *Bubble*-DNA, wiederholt werden. Dabei sollte vor allem die Datenqualität optimiert werden, um gute Ergebnisse bei einzelne Messungen zu erzielen. Weiterhin könnte die Affinitätsgruppe des HSAK+ *Crosslinkers*

verwendet werden, um *Crosslinker*-haltige Peptide anzureichern. Die zuvor durchgeführten *Crosslinking* Studien mit Anreicherung der *Crosslinks* durch SEC und die Auswertung mit einer modifizierten MaxQuant-Version zeigten fast keine *Crosslinks*.^[174] Außerdem wurde der Komplex nach Inkubation mit dem HSAK+ *Crosslinker* im SDS-Gel analysiert und ein In-Gel Verdau von einzelnen Banden, die vermutlich den Komplex enthielten, durchgeführt. Aber auch hier konnten fast keine *Crosslinks* erhalten werden. Erst der Wechsel der Messmethode (von pseudo-MS³ auf MS²) war für die ersten Ergebnisse ausschlaggebend. Auch die Optimierung der Inkubationstemperatur auf 0 °C und die Auswertung mit dem MeroX Programm zeigten etwas bessere Resultate. Zukünftig könnte aus den identifizierten Bindungsstellen mit Hilfe von Modellierungsprogrammen und den bisherigen Kristallstrukturen in Kooperation mit anderen Gruppen eine Struktur des Komplexes modelliert werden.

4.3.6. Kristallisationsstudien

Kristalle eines XPA/RPA/DNA-Komplexes können zur Aufklärung der dreidimensionalen Proteinstruktur *via* Röntgenstrukturanalyse beitragen, wodurch weitere Einblicke in den NER Mechanismus möglich werden.

Die Kristallisationsstudien für den XPA/RPA/DNA-Komplex wurden anfangs in enger Zusammenarbeit mit *Prof. Caroline Kisker* und *Dr. Jochen Kuper* (Würzburg) durchgeführt. Dabei ergaben die ersten Kristallisationsansätze mit den einzelnen Proteinen XPA und RPA keine Kristalle. Weiterhin wurde versucht, ein XPA/RPA-Komplex mit einem doppelsträngigen AAF-geschädigten 34mer (ODN: 34mer_1 und 34mer_3) zu kristallisieren. Dafür wurde der Protein/DNA-Komplex durch den Gelfiltrationsassay isoliert und auf ca. 1 mg/mL einkonzentriert. Dennoch konnten auch hier keine Kristalle nach der *sitting-drop* Methode bei 20 °C und bei verschiedenen Bedingungen erhalten werden. Nach Optimierung der DNA-Substrate für den XPA/RPA/DNA-Komplex wurde versucht doppelsträngige AAF-geschädigte DNA (34mer und 25mer) mit einem 20nt 5'-Überhang (ODN: 34mer_1 und 34mer_4, 25mer_1 und 25mer_4) bei 20 °C und bei verschiedenen Bedingungen zu kristallisieren. Lediglich konnten dabei nur paar Salzkristalle erhalten werden.

Am Ende dieser Arbeit wurde versucht die Kristallisationsbedingungen wie z.B. Temperatur und Proteinkonzentration zu optimieren, da die bisher durchgeführten Kristallisationsstudien zu keinem Erfolg führten. Dafür wurden verschiedene Kristallisationsansätze der einzelnen Proteine, sowie des XPA/RPA/DNA-Komplexes zusammen mit *Johanna Bretzler* nach der *sitting-drop* Methode bei 4 °C am MPI (Martinsried) angesetzt. Dabei wurde das AAF-geschädigte 25mer mit einem 20nt 5'-Überhang verwendet. Zudem wurde auch versucht die DBD von XPA (XPAs) mit einem AAF-geschädigten 15mer zu kristallisieren. Die Reinigung dieses Proteins ist bereits in Abschnitt 3.1. beschrieben. Tabelle 21 gibt einen Überblick der verschiedenen Kristallisationsansätze.

Tabelle 21: Überblick der verschiedenen Kristallisationsansätze. Beide DNAs enthielten einen AAF(C8)-dG Schaden.

Protein	XPAs	XPA	RPA	XPA/RPA/DNA-Komplex
DNA	15mer	25mer 20nt 5'-Überhang	25mer 20nt 5'-Überhang	25mer 20nt 5'-Überhang
C_{Protein}	1.) 3 mg/mL 2.) 4 mg/mL	1.) 3 mg/mL 2.) 4 mg/mL	1.) 7 mg/mL 2.) 8 mg/mL	1.) 1.64 mg/mL (XPA) und 5.50 mg/mL (RPA) 2.) 2.08 mg/mL (XPA) und 7 mg/mL (RPA)
Protein:DNA	2:1	2:1	2:1	2:2:1
Screens	1.) Qiagen Classic 2.) Qiagen Classic und MPD	1.) Qiagen Classic 2.) Qiagen Classic und MPD	1.) Qiagen Classic 2.) Qiagen Classic	1.) Qiagen Classic und PACT, Molecular Dimension JCSG+ 2.) Qiagen Classic, MPD und Nucleix
Temperatur	4 °C	4 °C	4 °C	4 °C

Auch in diesen Kristallisationsstudien konnten keine Kristalle erhalten werden. Dafür konnten aber Mikrokristalle für den XPA/RPA/DNA-Komplex nach einem Tag (0.2 M MgCl₂, 0.1 M HEPES·NaCl pH 7.5, 30% (v/v) Isopropanol) erhalten werden. Diese liefern erste Hinweise für Kristallisationsbedingungen des Komplexes. In Zukunft werden die Kristallisationsstudien von *Johanna Bretzler* weitergeführt.

4.3.7. Kryo-Elektronenmikroskopie

Ein anderer Fokus zur Strukturlösung des XPA/RPA/DNA-Komplexes ist die Kryo-EM. Dabei wurden vorweg Negativfärbung Studien durchgeführt, um zum einen die Homogenität und die Partikelgröße zu bestimmen und zum anderen um die optimale Komplexkonzentration für Kryo-EM Studien zu finden. Hierbei wurden Negativfärbung Experimente in Zusammenarbeit mit der Beckmann Gruppe (LMU, München) von mir durchgeführt. Für diese Studien wurde ein doppelsträngiges AAF-geschädigtes 34mer mit einem 20nt 5'-Überhang (ODN: 34mer_1 und 34mer_4) verwendet. Der durch den Gelfiltrationsassay isolierte XPA/RPA/DNA-Komplex wurde zur Konzentrationsbestimmung einkonzentriert (ca. 0.3-0.7 mg/mL), da dieser ansonsten zur Bestimmung der Konzentration zu verdünnt vorlag. Anschließend wurden verschiedene Verdünnungen für die Negativfärbung Studien getestet (ca. 44-150 µg/mL). Erste Ergebnisse zeigten zwar teils Partikel, aber diese waren sehr inhomogen und teilweise aggregiert. Dennoch konnten erste Anhaltspunkte für eine optimale Verdünnung erhalten werden. Außerdem wurde der isolierte Komplex mit Glutaraldehyd vor der Färbung für 10 min inkubiert, um homogenere Partikel zu erhalten. Allerdings führte dies auch zu keiner Verbesserung.

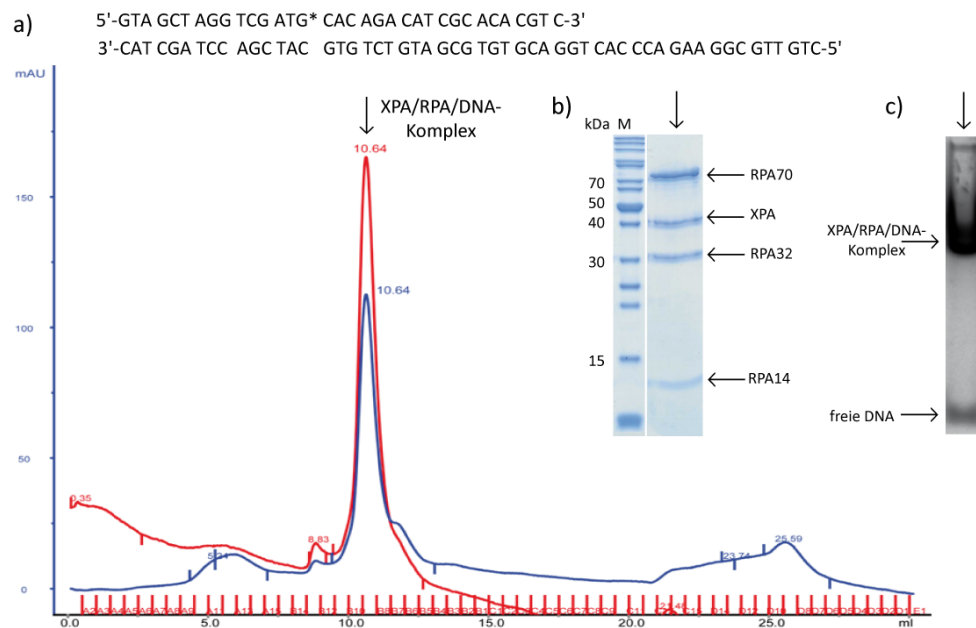


Abbildung 41: a) Chromatogramm des Gelfiltrationsassays und Sequenz der verwendeten DNA. G*: AAF(C8)-dG. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. Der XPA/RPA/DNA-Komplex hat ein Retentionsvolumen von 10.64 mL. **b) SDS-PAGE der Komplex-enthaltende Fraktion.** M: Proteinstandard PageRuler™, XPA und RPA liegen in einem 1:1 Verhältnis vor. **c) Natives Gel der Komplex-enthaltenden Fraktion.** Die DNA wird von den Proteinen gebunden.

Nach diesen ersten Vormessungen am Morgagni-Mikroskop wurde der Komplex erneut isoliert (Abbildung 41) und auf ca. 0.48 mg/mL einkonzentriert. Durch die zuvor gewonnenen Kenntnisse

wurde der Komplex 1:10 (48 µg/mL) und 1:20 (24 µg/mL) verdünnt. Diese Verdünnungen zeigten bessere Resultate (Abbildung 42a und 42b), sodass eine Datensammlungen von mehreren Grids durch *Charlotte Ungewickel* am Spirit-Mikroskop (Beckmann Gruppe, LMU, München) erfolgte.

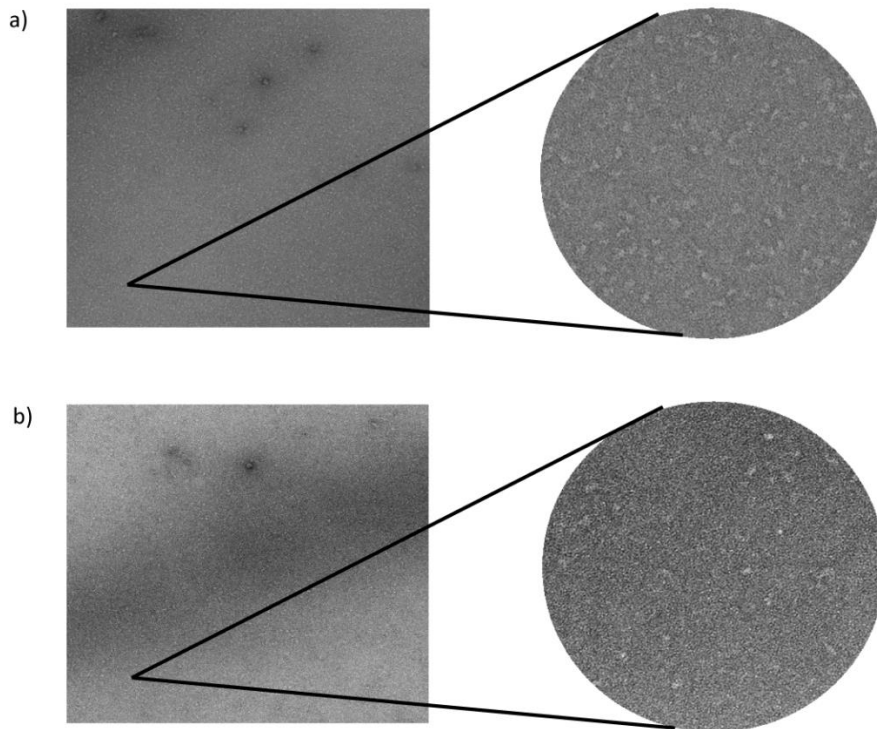


Abbildung 42: Ausschnitte der Negativfärbung Grids (Mikrographen) mit Zoom von der Verdünnung a) 1:10 (48 µg/mL) und b) 1:20 (24 µg/mL). Auflösung der Mikrographen: 1.28 µM x 1.28 µM (4096 x 4096 Pixel), 3.2 pixel/nm. Die Datensammlung der einzelnen Mikrographen erfolgte durch *Charlotte Ungewickel* am Spirit-Mikroskop (Beckmann Gruppe, LMU, München).

Bei beiden Verdünnungen liegen sichtbare und größtenteils homogene Partikel vor. Bei dem Grid mit der 1:20 Verdünnung (Abbildung 42b) war die Färbung nicht konstant, wodurch teilweise ein hell-dunkel Gradient entstanden ist. Außerdem waren nur sehr wenige Partikel vorhanden. Im Vergleich dazu, lassen sich bei der 1:10 Verdünnung (Abbildung 42a) mehr Partikel finden. Teilweise sogar zu viele, wodurch manche Partikel sich gegenseitig berühren. Wahrscheinlich wäre eine Verdünnung von 1:15 optimal gewesen. Auch bei dem Grid mit der 1:10 Verdünnung war die Färbung teils nicht homogen, aber dennoch besser als bei der 1:20 Verdünnung. Hinzu lassen sich noch einzelne Aggregate bei beiden Verdünnungen erkennen. Dennoch sind die Partikel homogen genug, um eine erste 2D-Klassifizierung zur Bestimmung der Partikelgröße und -form durchführen zu können. Aufgrund der vorher genannten Faktoren wurden die Partikel für die 2D-Klassifizierung manuell ausgewählt. Die 2D-Klassifikation der Partikel wurde von *Johanna Bretzler* unter Anleitung von *Sebastian Eustermann* (Hopfner Gruppe, LMU) durchgeführt. Bei der Evaluierung konnten für beide

Verdünnungen unabhängig voneinander ähnliche Partikelklassen identifiziert werden. Im Folgenden wird die 2D-Klassifizierung der 1:10 Verdünnung veranschaulicht (Abbildung 43).

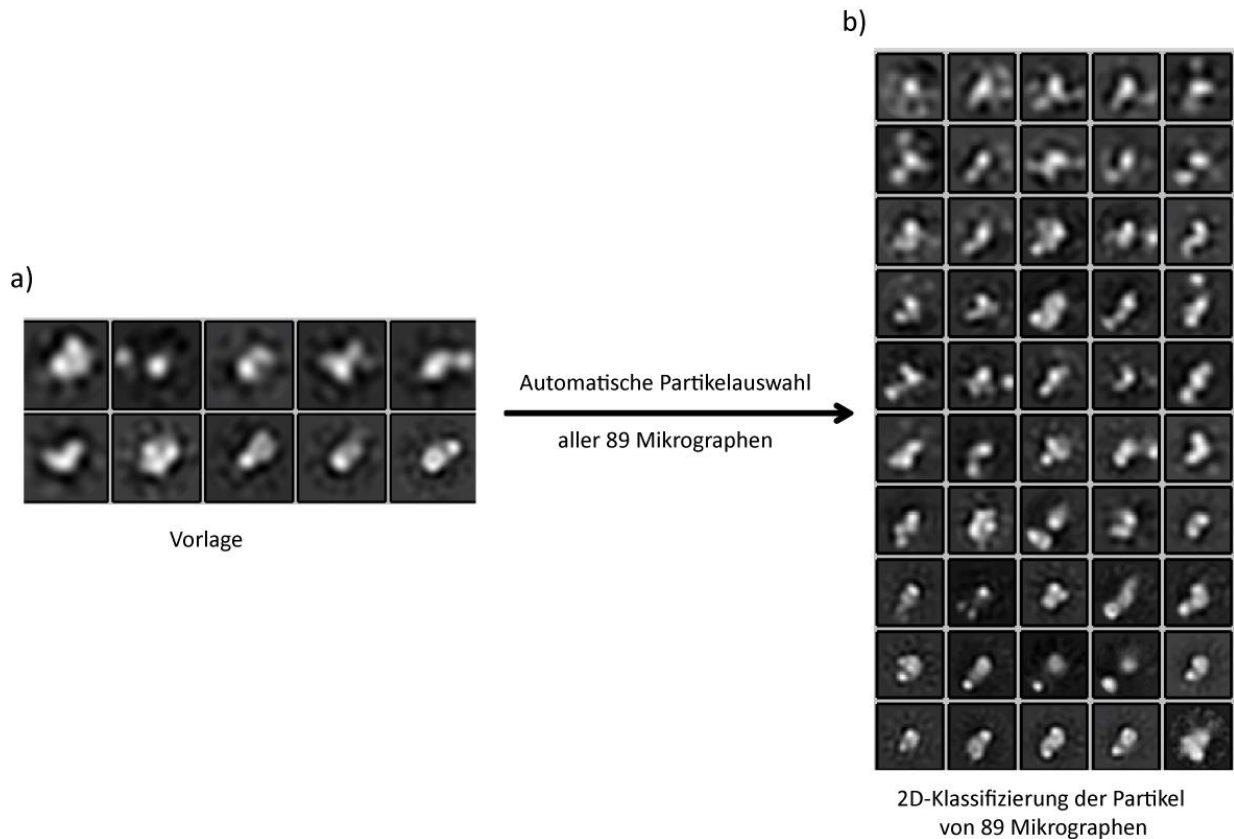


Abbildung 43: 2D-Klassifizierung der Partikel von der 1:10 Verdünnung. a) 2D-Klassifizierung und Erhalt von 10 Klassen durch das Programm EMAN2. Dabei wurden 1135 Partikel in 6 Mikrographen manuell ausgewählt. b) 2D-Klassifizierung und Erhalt von 50 Klassen durch das Programm Relion-1.3. Dabei dienten die 10 Klassen als *Template*, wodurch 92764 Partikel automatisch in 89 Mikrographen ausgewählt wurden. Durchführung der 2D-Klassifikation erfolgte durch *Johanna Bretzler* unter Anleitung von *Sebastian Eustermann* (Hopfner Gruppe, LMU) durchgeführt.

Im ersten Schritt der Evaluierung wurden insgesamt 1135 Partikel in 6 Mikrographen manuell ausgewählt und einer 2D-Klassifizierung unterzogen. Dabei konnten 10 Klassen erhalten werden, welche als Vorlage für die automatische Partikelauswahl fungieren. Dabei wurden 92764 Partikel automatisch ausgewählt, darunter auch viele falsche Partikel. Daher musste eine manuelle Deselektion falscher Partikel erfolgen. Letztendlich wurde dann mit 40000 Partikel eine 2D-Klassifizierung durchgeführt. Dabei konnten 50 Klassen identifiziert werden, die aber teilweise Artefakte wie Schmier oder kleine weiße Punkte aufweisen. Diese könnten durch die unregelmäßige Färbung, durch falsche Partikel, durch nicht zentrierte Partikel und durch die schlechte Auflösung der Datensammlung verursacht werden. Somit sind die Ergebnisse der 2D-Klassifizierung nicht ausreichend für eine weitere 3D-Klassifizierung. Homogenere Präparationen und eine bessere Auflösung der Klassen bzw. Partikel sind notwendig. Außerdem wäre ein Strukturmodell des Komplexes als Vorlage für die 3D-Klassifizierung sehr hilfreich.

Dennoch sind die ersten Ergebnisse vielversprechend. Denn es sind Partikel erkennbar, die richtige Konzentration des Komplexes konnte gefunden werden und die Evaluierung beider Verdünnungen zeigten unabhängig voneinander die gleichen Klassen in der 2D-Klassifizierung. Dadurch konnte eine Partikelgröße des Komplexes von ca. 6-15 nm errechnet werden. Zudem wird bei den einigen Klassen erkennbar, dass die Partikel bzw. der Komplex eine Art Kirschstruktur einnimmt. In Kryo-EM Studien von *Treuner et al.* wurde RPA mit Glutaraldehyd an die DNA fixiert und mittels Gelfiltration ungebundenes Protein abgetrennt.^[109] Auch *Witosch et al.* benutzten in Kryo-EM Studien Glutaraldehyd zur Fixierung des Tim-Tipin/RPA-Komplexes an die ssDNA.^[111] Die Verwendung von Glutaraldehyd für den XPA/RPA-Komplex zur Fixierung an die DNA und eine anschließende Gelfiltration wäre eine interessante Überlegung für zukünftige Experimente, um die Homogenität der Partikel weiter zu optimieren. Ein weiterer interessanter Punkt ist die Konformation der Kryo-EM Struktur des Tim-Tipin/RPA/ssDNA-Komplexes von *Witosch et al.*^[111] In dieser nimmt RPA eine hufeneisenförmige Konformation entlang der ssDNA ein. Zudem zeigt diese U-Konformation eine gewisse Ähnlichkeit mit der zuvor durch 2D-Klassifizierung erhaltenen Kirschkonformation des XPA/RPA/DNA-Komplexes auf. Dabei könnte der Knick in der Kirschkonformation durch XPA verursacht worden sein. Denn bereits in Abschnitt 3.3. konnte gezeigt werden, dass XPA die DNA knickt. Durch diese Ähnlichkeit beider Strukturen könnte die Kryo-EM Struktur von RPA mit ssDNA im Tim-Tipin/RPA/ssDNA-Komplex für spätere Evaluierungen des XPA/RPA/DNA-Komplexes als Vorlage fungieren.

Nach Erhalt der ersten Partikel und die optimale Konzentration des Komplexes in den Negativfärbung Studien wurden Kryo-EM Messungen für den Komplex in Kooperation mit der Dietz Gruppe (TU München) durchgeführt. Für diese Studien wurde ein doppelsträngiges AAF-geschädigtes 25mer mit einem 20nt 5'-Überhang (ODN: 25mer_1 und 25mer_4) verwendet. Auch hier wurde der XPA/RPA/DNA-Komplex durch den Gelfiltrationsassay isoliert. Leider waren bis zum Ende dieser Arbeit keine Partikel in den Kryo-EM Studien erkennbar. Vermutlich müssen die Einstellung des Mikroskops weiter optimiert werden oder es könnte versucht werden größere Partikel zu erhalten. Dabei könnte die Isolierung eines größeren Komplexes mit einem weiteren NER-Protein die Bildung von größeren Partikeln ermöglichen.

4.3.8. Zusammenfassung

In diesem Abschnitt konnten neben der Synthese von verschiedenen Oligonukleotiden mit einem C8-AAF-dG Schaden die Proteine XPA und RPA erfolgreich exprimiert und aufgereinigt werden. Weiterhin wurde die Etablierung eines neuen Gelfiltrationsassays, der durch verschiedene Bedingungen effektiv optimiert werden konnte, vorgestellt. Dieser ermöglichte die Isolierung eines stabilen XPA/RPA/DNA-Komplexes. Zusätzlich konnten durch den Gelfiltrationsassay verschiedene DNA-Substrate für die Komplexbildung getestet werden, um das am besten geeignete Substrat zu finden. Zusammenfassend konnten folgende DNA-Substrate für eine sehr gute Komplexbildung erhalten werden: Die ungeschädigten, als auch die AAF-geschädigten ds DNA-Stränge mit Überhang (ds 34mer mit 20nt 5'-Überhang und ds 25mer mit 20nt 5'-Überhang) und verschiedene lange AAF-geschädigte *Bubble*-DNAs. Da sowohl die ungeschädigte als auch die geschädigte DNA mit Überhang einen Komplex ergaben, erkennt der XPA/RPA-Komplex wahrscheinlich die ss/ds-Verzweigung der DNA, also nicht AAF, als Schaden. Im Gegensatz dazu ist die Komplexbildung mit *Bubble*-DNAs schadensabhängig. Diese DNA-Substrate konnten dann für verschiedene Strukturstudien verwendet werden. Ferner ergaben Kristallisationsstudien des XPA/RPA/DNA-Komplexes mit verschiedenen DNA-Substraten zwar keine Bildung von Kristallen, dafür aber Mikrokristalle. Diese könnten erste gute Hinweise auf die Kristallisationsbedingungen des Komplexes geben. Gleichzeitig wurde versucht, die Struktur des Komplexes durch Kryo-Elektronenmikroskopie zu lösen. Dabei konnte durch erste Negativfärbungen eine Partikelgröße von ca. 6-15 nm und durch 2D-Klassifizierung eine Art Kirschstruktur des XPA/RPA/DNA-Komplexes (DNA-Substrat: ds 34mer AAF mit 20nt 5'-Überhang) erhalten werden. Dennoch konnte bisher keine hochaufgelöste Struktur mittels Kryo-EM gelöst werden. In zukünftigen Arbeiten könnte der Gelfiltrationsassay für die Komplexbildung von XPA/RPA/DNA genutzt werden, um weitere Strukturstudien durchzuführen.

5. Ausblick

Die verschiedenen Kristallstrukturen von Rad14 mit unterschiedlichen DNA-Schäden erbrachten einen ersten Einblick in die Schadenserkennung bzw. -verifizierung von XPA/Rad14 in der NER. Darüber hinaus wäre es interessant, weitere Kristallstrukturen mit längerer oder mit einer größeren *Bubble*-DNA zu erhalten, da diese dem natürlichen NER-Substrat näher kommen. Um den Mechanismus von Rad14 weiter zu verifizieren, ist eine Kristallstruktur des humanen Proteins XPA mit DNA von großer Bedeutung.

Von hohem Interesse sind außerdem die strukturelle Interaktion von XPA mit anderen Proteinen und die Frage, wie XPA die einzelnen Proteine im NER-Komplex positioniert. Dies könnte vor allem zeigen, ob XPA als Gerüstprotein fungieren kann, und weitere Erkenntnisse im NER Mechanismus bringen. Besonders die strukturelle Aufklärung des XPA/RPA/DNA-Komplexes könnte mechanistische Details für den NER-Mechanismus liefern. Dafür kann der in dieser Arbeit entwickelte Gelfiltrationsassay zur Isolierung des Komplexes weiter verwendet werden. Weiterhin müssen die Kristallisationsbedingungen vom isolierten Komplex optimiert werden, um Kristalle und somit eine mögliche Struktur zu erhalten. Daneben könnte auch versucht werden, nur die DBD von XPA mit RPA70ABC/RPA32D und DNA zu kristallisieren, um die flexiblen Domänen von beiden Proteinen, die die Kristallisation erschweren, zu vermeiden. Zusätzlich könnten auch verschiedene DNA-Substrate, wie z.B. eine größere *Bubble* im DNA-Strang, für die Kristallisation getestet werden.

Ferner könnte zur Strukturaufklärung auch die Bedingungen für die Kryo-Elektronenmikroskopie weiter optimiert werden. Dafür müsste versucht werden, homogenere Partikel zu erhalten. Hinzu könnten auch andere DNA-Substrate, wie z.B. die *Bubble*-DNA, getestet werden. Ebenso könnte ein DNA-Protein-Crosslinker oder Glutaraldehyd verwendet werden, der mit dem Komplex reagieren und somit die Flexibilität der Proteine reduzieren kann. Anschließend kann dieser über den Gelfiltrationsassay aufgereinigt und isoliert werden. Dies könnte eine bessere Homogenität der Partikel für die Kryo-EM erzeugen. Daneben könnte auch versucht werden einen Komplex aus XPA/RPA/DNA mit einem weiteren NER-Protein zu isolieren. Dieses würde die Ausbildung eines größeren Komplexes ermöglichen, woraufhin eine leichtere Strukturauflösung *via* Kryo-EM möglich wäre.

Proteomik-Daten könnten Erkenntnisse in der funktionellen und strukturellen Beziehung der einzelnen Proteine liefern. Vor allem weitere Untersuchungen von XPA/RPA mit einem Protein-Protein-Crosslinker mit und ohne DNA könnten Indizien für den NER-Mechanismus, insbesondere ob die Interaktion von RPA über einen allosterischen Mechanismus verläuft, liefern. Aus den identifizierten Bindungsstellen mittels HPLC-MS/MS könnte wahrscheinlich mit Hilfe von

Modellierungsprogrammen eine Struktur des XPA/RPA/DNA-Komplexes modelliert werden. Außerdem könnte auch ein DNA-Protein-*Crosslinker* verwendet werden, wodurch die genaue Lokalisation der Proteine an der DNA durch Analyse mit HPLC-MS/MS ermittelt werden könnte.

Das Verstehen des NER-Mechanismus würde in Zukunft wichtige Aspekte in der Medikamentenentwicklung ermöglichen. Denn das NER-System kann die Resistenz von Chemotherapeutika, wie z.B. Cisplatin, fördern, indem es die induzierten DNA-Quervernetzungen repariert. Die Entwicklung von XPA Inhibitoren könnte dem entgegenwirken.

6. Anhang

6.1. Gensequenzen XPA

cDNA-Sequenz des *xpaf1*-Gens aus *H. sapiens* (819 bp)

ATGGCGGCGGCCGACGGGGCTTTGCCGGAGGCGGCGGCTTTAGAGCAACCCGCGGAGCTGCCTGCCTCGGT
GCGGGCGAGTATCGAGCGGAAGCGGCAGCGGGCACTGATGCTGCGCCAGGCCCGGCTGCTGCCCCGGCCCTA
CTCGGCGACGGCGGCTGCGGCTACTGGAGGCATGGCTAATGTAAAAGCAGCCCCAAAGATAATTGACACAG
GAGGAGGCTTCATTTAGAAGAGGAAGAAGAAGAAGAACAGAAAATTGGAAAAGTTGTTTCATCAACCAGGA
CCTGTTATGGAATTTGATTATGTAATATGCGAAGAATGTGGGAAAGAATTTATGGATTCTTATCTTATGAACCA
CTTTGATTTGCCAACTTGTGATAACTGCAGAGATGCTGATGATAAACACAAGCTTATAACCAAAACAGAGGCA
AAACAAGAATATCTTCTGAAAGACTGTGATTTAGAAAAAAGAGAGCCACCTCTTAAATTTATTGTGAAGAAGA
ATCCACATCATTACAATGGGGTGATATGAACTCTACTTAAAGTTACAGATTGTGAAGAGGTCTCTTGAAGTT
TGGGGTAGTCAAGAAGCATTAGAAGAAGCAAAGGAAGTCCGACAGGAAAACCGAGAAAAAATGAAACAGA
AGAAATTTGATAAAAAAGTAAAGAATTGCGGCGAGCAGTAAGAAGCAGCGTGTGAAAAGGGAGACGATT
GTTTCATCAACATGAGTATGGACCAGAAGAAAACCTAGAAGATGACATGTACCGTAAGACTTGTACTATGTGTG
GCCATGAACTGACATATGAAAAAATGTAA

cDNA-Sequenz des *xpas*-Gens aus *H. sapiens* (DNA-bindendes Fragment, 366 bp)

ATGGAATTTGATTATGTAATATGCGAAGAATGTGGGAAAGAATTTATGGATTCTTATCTTATGAACCACTTTGA
TTTGCCAACTTGTGATAACTGCAGAGATGCTGATGATAAACACAAGCTTATAACCAAAACAGAGGCAAAACAA
GAATATCTTCTGAAAGACTGTGATTTAGAAAAAAGAGAGCCACCTCTTAAATTTATTGTGAAGAAGAATCCAC
ATCATTACAATGGGGTGATATGAACTCTACTTAAAGTTACAGATTGTGAAGAGGTCTCTTGAAGTTTGGGG
TAGTCAAGAAGCATTAGAAGAAGCAAAGGAAGTCCGACAGGAAAACCGAGAAAAAATGAAACAGAAGAAAT
TT

6.2. Gensequenzen Rad14

cDNA-Sequenz des *rad14s*-Gens aus *S. cerevisiae* (DNA-bindendes Fragment, 363 bp)

ATGGCGCCGAAATGTATTGAATGTCATATTAATATTGAGATGGATCCTGTGCTACATGATGTGTTCAAGTTACA
AGTTTGTAACAGTGTTCTAAGGAGCATCCAGAAAAGTATGCACTACTGACGAAAACAGAATGTAAGGAAGA
TTACTTTTTAACAGACCCCGAATTGAATGATGAGGATCTCTTTCATAGACTAGAAAAGCCGAACCCCTCATTCGG
GGACATTTGCAAGAATGCAACTATTTGTTAGATGTGAAGTGAAGCCTTTCGTTCAAGAAATGGGGTGGAG
AAGAAGGTTTAGATGAGGAATGGCAACGTCGTGAAGAAGGAAAGGCTCACAGAAGGGAGAAAAAATACTAA

6.3. Gensequenzen RPA

cDNA-Sequenz des *rpa14*-Gens aus *H. Sapiens* (RPA14-Untereinheit, 363 bp)

ATGGTGGACATGATGGACTTGCCAGGTCGCGCATCAACGCCGGCATGCTAGCTCAATTCATCGACAAGCCTG
TCTGCTTCGTAGGGAGGCTGGAAAAGATTCATCCCACCGGAAAAATGTTTATTCTTTCAGATGGAGAAGGAAA
AAATGGAACCATCGAGTTGATGGAACCCCTTGATGAAGAAATCTCTGGAATTGTGGAAGTGGTTGGAAGAGT
AACCGCCAAGGCCACCATCTTGTGTACATCTTATGTCCAGTTTAAAGAAGATAGCCATCCTTTTGATCTTGGAC
TTTACAATGAAGCTGTGAAAATTATCCATGACTTCCCTCAGTTTTATCCTTTAGGGATTGTGCAACATGAT

cDNA-Sequenz des *rpa32*-Gens aus *H. sapiens* (RPA32-Untereinheit, 810 bp)

ATGTGGAACAGTGGATTCGAAAGCTATGGCAGCTCCTCATACGGGGGAGCCGGCGGCTACACGCAGTCCCCG
GGGGGCTTTGGATCGCCCGCACCTTCTCAAGCCGAAAAGAAATCAAGAGCCCGAGCCAGCACATTGTGCCC
TGTAATATATCTCAGCTGCTTTCTGCCACTTTGGTTGATGAAGTGTTGAGAATTGGGAATGTTGAGATTTACA
GGTCACTATTGTGGGGATCATCAGACATGCAGAGAAGGCTCCAACCAACATTGTTTACAAAATAGATGACATG
ACAGCTGCACCCATGGACGTTCCGAGTGGGTTGACACAGATGACACCAGCAGTGAAAACACTGTGGTTCCTC
CAGAAACATATGTGAAAGTGGCAGGCCACCTGAGATCTTTTCAGAACAAAAAGAGCCTGGTAGCCTTTAAGA
TCATGCCCCTGGAGGATATGAATGAGTTCACCACACATATTCTGGAAGTGATCAATGCACACATGGTACTAAG
CAAAGCCAACAGCCAGCCCTCAGCAGGGAGAGCACCTATCAGCAATCCAGGAATGAGTGAAGCAGGGAACTT
TGGTGGGAATAGCTTCATGCCAGCAAATGGCCTCACTGTGGCCCAAAACCAGGTGTTGAATTTGATTAAGGCT
TGTCCAAGACCTGAAGGGTTGAACTTTAGGATCTCAAGAACCAGCTGAAACACATGTCTGTATCCTCAATCA
AGCAAGCTGTGGATTTTCTGAGCAATGAGGGGCACATCTATTCTACTGTGGATGATGACCATTTTAAATCCAC
AGATGCAGAA

cDNA-Sequenz des *rpa70*-Gens aus *H. sapiens* (RPA70-Untereinheit, 1.848 bp)

ATGGTCGGCCAGCTGAGCGAGGGGGCCATTGCGGCCATCATGCAGAAGGGGGATACAAACATAAAGCCCAT
CCTCCAAGTCATCAACATCCGTCCCATTAACGCGGGAATAGTCCGCCGCGTTATCGACTGCTCATGAGTGATG
GATTGAACACTCTATCCTCTTTTCATGTTGGCGACACAGTTGAACCCTCTCGTGGAGGAAGAACAATTGTCCAGC
AACTGTGTATGCCAGATTCACAGATTTATTGTGAACACTCTGAAAGACGGAAGGAGAGTAGTTATCTTGATGG
AATTAGAAGTTTTGAAGTCAGCTGAAGCAGTTGGAGTGAAGATTGGCAATCCAGTGCCCTATAATGAAGGAC
TCGGGCAGCCGCAAGTAGCTCCTCCAGCGCCAGCAGCCAGCCAGCAGCAAGCAGCAGGCCCCAGCCGCAG
AATGGAAGCTCGGGAATGGGTTCTACTGTTTCTAAGGCTTATGGTGCTTCAAAGACATTTGGAAAAGCTGCAG
GTCCAGCCTGTCACACACTTCTGGGGGAACACAGTCCAAAGTGGTGCCCATTGCCAGCCTCACTCCTTACCA
GTCCAAGTGGACCATTTGTGCTCGTGTACCAACAAAAGTCAGATCCGTACCTGGAGCAACTCCCGAGGGGAA
GGGAAGCTTTTCTCCCTAGAACTGGTTGACGAAAGTGGTGAAATCCGAGCTACAGCTTTCAATGAGCAAGTG
GACAAGTTCTTTCTTATTGAAGTGAACAAGGTGTATTATTTCTCGAAAGGCACCCTGAAGATTGCTAACAA
GCAGTTCACAGCTGTAAAAATGACTACGAGATGACCTTCAATAACGAGACTTCCGTTCATGCCCTGTGAGGAC
GACCATCATTTACCTACGGTTCAGTTTGATTTACGGGGATTGATGACCTCGAGAACAAGTCGAAAGACTCAC
TTGTAGACATCATCGGGATCTGCAAGAGCTATGAAGACGCCACTAAAATCACAGTGAGGTCTAACACAGAG
AAGTTGCCAAGAGGAATATCTACTTGATGGACACATCCGGGAAGGTGGTGACTGCTACACTGTGGGGGGA
AGATGCTGATAAATTTGATGGTTCTAGACAGCCCGTGTGGCTATCAAAGGAGCCCGAGTCTCTGATTTCCGT
GGACGGAGCCTCTCCGTGCTGTCTCAAGCACTATCATTGCGAATCCTGACATCCCAGAGGCCTATAAGCTTCG
TGGATGGTTTGACGCAGAAGGACAAGCCTTAGATGGTGTTCATCTCTGATCTAAAGAGCGGCGGAGTCGG
AGGGAGTAACACCAACTGGAAAACCTTGATGAGGTCAAATCCGAGAACCTGGGCCAAGGCGACAAGCCGG
ACTACTTTAGTTCTGTGGCCACAGTGGTGTATCTTCGCAAAGAGAACTGCATGTACCAAGCCTGCCGACTCA
GGACTGCAATAAGAAAGTGATTGATCAACAGAATGGATTGTACCGCTGTGAGAAGTGCGACACCGAATTTCC
CAATTTCAAGTACCGCATGATCCTGTGAGTAAATATTGCAGATTTTCAAGAGAATCAGTGGGTGACTTGTTTCC
AGGAGTCTGCTGAAGCTATCCTTGGACAAAATGCTGCTTATCTTGGGGAATTAAGACAAGAATGAACAGG
CATTTGAAGAAGTTTTCCAGAATGCCAACTCCGATCTTTCATATTCAGAGTCAGGGTCAAAGTGGAGACCTAC
AACGACGAGTCTCGAATTAAGGCCACTGTGATGGACGTGAAGCCCGTGGACTACAGAGAGTATGGCCGAAG
GCTGGTCATGAGCATCAGGAGAAGTGCATTGATG

6.4. Pufferbedingungen für den Thermofluor Assay

Tabelle 22: Pufferbedingungen für den Thermofluor Assay nach Boivin et al.^[162] Die Pufferkonzentration betrug 100 mM, außer anders angegeben. Die Natriumchlorid Konzentration betrug 250 mM, außer anders angegeben. Puffer A Zusammensetzung: Bernsteinsäure/NaPO₄/Glycin [2:7:7]; Puffer B Zusammensetzung: Zitronensäure/CHES/HEPES [2:4:3].

	1	2	3	4	5	6
A	H ₂ O	Zitrone-säure pH 4.0	Natriumacetat pH 4.5	Zitronensäure pH 5.0	MES pH 6.0	KH ₂ PO ₄ pH 6.0
B	MOPS pH 7.0	Ammonium- acetat pH 3	Tris-HCl pH 7.5	NaH ₂ PO ₄ pH 7.5	Imidazol pH 8.0	HEPES pH 8.0
C	H ₂ O NaCl	Zitronensäure, NaCl pH 4.0	Natrium- acetat, NaCl pH 4.5	Zitronen- säure, NaCl pH 5.0	MES NaCl pH 6.0	KH ₂ PO ₄ NaCl pH 6.0
D	MOPS NaCl pH 7.0	Ammonium- acetat, NaCl pH 3	Tris-HCl NaCl pH 7.5	NaH ₂ PO ₄ NaCl pH 7.5	Imidazol NaCl pH 8.0	HEPES NaCl pH 8.0
E	Puffer A pH 4.0	Puffer A pH 4.78	Puffer A pH 5.21	Puffer A pH 5.62	Puffer A pH 5.95	Puffer A pH 6.23
F	Puffer B pH 4.0	Puffer B pH 4.22	Puffer B pH 4.62	Puffer B pH 5.06	Puffer B pH 5.62	Puffer B pH 6.49
G	10 mM HEPES pH 7.5	50 mM HEPES pH 7.5	100 mM HEPES pH 7.5	250 mM HEPES pH 7.5	10 mM NaPO ₄ pH 7.5	50 mM NaPO ₄ pH 7.5
H	50 mM HEPES 50 mM NaCl pH 7.5	50 mM HEPES 125 mM NaCl pH 7.5	50 mM HEPES 250 mM NaCl pH 7.5	50 mM HEPES 500 mM NaCl pH 7.5	50 mM HEPES 750 mM NaCl pH 7.5	50 mM HEPES 1 M NaCl pH 7.5
	7	8	9	10	11	12
A	Zitronensäure pH 7.0	Bis- Tris pH 6.5	Natriumkako- dylat pH 6.5	NaH ₂ PO ₄ pH 7.0	KH ₂ PO ₄ pH 7.0	HEPES pH 7.0
B	Tris-HCl pH 8.0	Tricin pH 8.0	Bicin pH 8.0	Bicin pH 8.5	Tris-HCl pH 8.5	Bicin pH 9.0
C	Zitronen- säure, NaCl pH 7.0	Bis-Tris NaCl pH 6.5	Natriumkako- dylat, NaCl pH 6.5	Zitronen- säure, NaCl pH 7.0	Bis-Tris NaCl pH 6.5	Natriumkako- dylat, NaCl pH 6.5
D	Tris-HCl NaCl pH 8.0	Tricin, NaCl pH 8.0	Bicin, NaCl pH 8.0	Tris-HCl NaCl pH 8.0	Tricin, NaCl pH 8.0	Bicin, NaCl pH 8.0
E	Puffer A pH 6.53	Puffer A pH 6.81	Puffer A pH 7.16	Puffer A pH 7.80	Puffer A pH 9.0	Puffer A pH 10.0
F	Puffer B pH 7.25	Puffer B pH 7.76	Puffer B pH 8.20	Puffer B pH 8.75	Puffer B pH 9.0	Puffer B pH 10.0
G	100 mM NaPO ₄ pH 7.5	200 mM NaPO ₄ pH 7.5	10 mM Tris-HCl pH 8.0	50 mM Tris- HCl pH 8.0	100 mM Tris- HCl pH 8.0	250 mM Tris-HCl pH 8.0
H	50 mM Tris-HCl 50 mM NaCl pH 8.0	50 mM Tris-HCl 125 mM NaCl pH 8.0	50 mM Tris-HCl 250 mM NaCl pH 8.0	50 mM Tris-HCl 500 mM NaCl pH 8.0	50 mM Tris-HCl 750 mM NaCl pH 8.0	50 mM Tris-HCl 1 M NaCl pH 8.0

Tabelle 23: Selbst getestete Pufferbedingungen für den Thermofluor Assay. Xtal I: Kristallisationspuffer I; HepA: Heparin Puffer A von XPA. Die Screens wurden mit den Proteinen alleine getestet, XPA und RPA zusammen sowie DNA mit Protein.

	1	2	3	4	5	6	7
A	Xtal I, 0% Glycerin	Xtal I, 1% Glycerin	Xtal I, 2% Glycerin	Xtal I, 3% Glycerin	Xtal I, 4% Glycerin	Xtal I, 5% Glycerin	HepA

6.5. Gelfiltrationsassay

6.5.1. Retentionsvolumina der einzelnen DNA-Sequenzen

Tabelle 24: Retentionsvolumina der verwendeten DNA-Sequenzen *via* Gelfiltration.

ds DNA Substrat	Retentionsvolumina [mL]	DNA-Schaden
15mer	17.3	Ungeschädigt
15mer	17.3	AAF(C8)-dG
15mer mit ss 20nt 5'-Überhang	15.92	Ungeschädigt
15mer mit ss 20nt 5'-Überhang	15.57	AAF(C8)-dG
25mer	15.90	Ungeschädigt
25mer	15.71	AAF(C8)-dG
25mer mit ss 20nt 5'-Überhang	14.64	Ungeschädigt
25mer mit ss 20nt 5'-Überhang	14.64	AAF(C8)-dG
25mer mit ss 15nt 5'-Überhang	14.96	AAF(C8)-dG
25mer mit ss 10nt 5'-Überhang	15.08	AAF(C8)-dG
25mer mit 3 <i>Mismatches</i> an beiden Enden	15.85	AAF(C8)-dG
34mer	14.68	AAF(C8)-dG
34mer mit ss 20nt 5'-Überhang	14.10	Ungeschädigt
34mer mit ss 20nt 5'-Überhang	13.71	AAF(C8)-dG
34mer mit ss 15nt 5'-Überhang	14.31	AAF(C8)-dG
34mer mit ss 10nt 5'-Überhang	14.24	AAF(C8)-dG
33mer mit 13mer <i>Bubble</i>	15.31	Ungeschädigt
33mer mit 13mer <i>Bubble</i>	15.27	AAF(C8)-dG
34mer mit 13mer <i>Bubble</i>	14.84	AAF(C8)-dG
39mer mit 16mer <i>Bubble</i>	14.71	Ungeschädigt
39mer mit 16mer <i>Bubble</i>	14.99	AAF(C8)-dG
45mer mit 25mer <i>Bubble</i>	14.19	Ungeschädigt

6.5.2. Gelfiltrationsassay mit den einzelnen Proteinen und DNA

Die Proteine XPA und RPA wurden einzeln mit einigen DNA-Substraten mittels Gelfiltration analysiert, um zu testen, ob diese auch alleine binden können. Dabei konnten XPA und RPA an alle, die hier zeigten, DNA-Substrate binden.

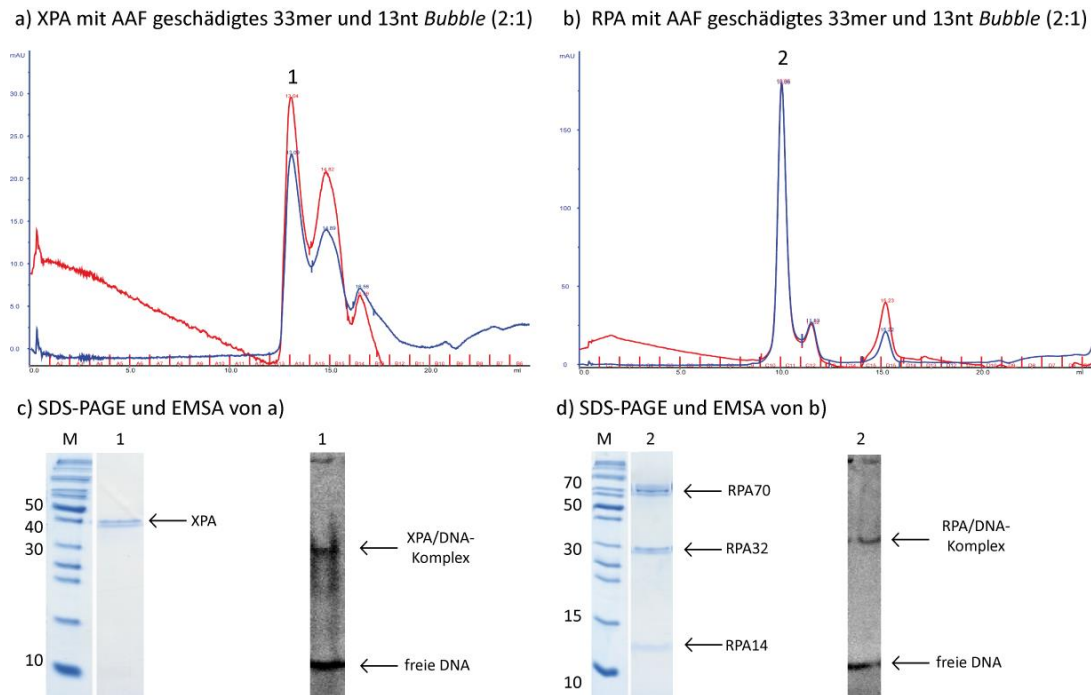
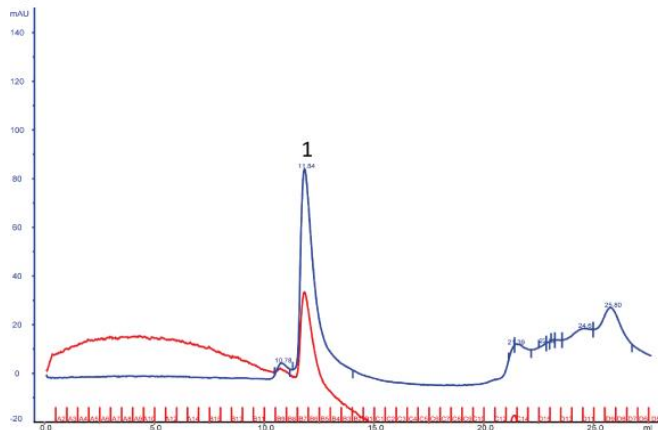
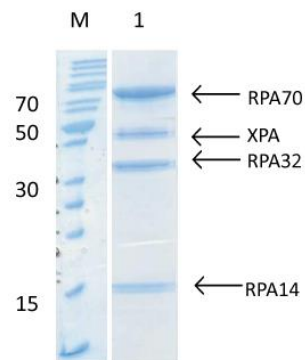


Abbildung 44: a) XPA mit AAF geschädigtem 33mer und einer 13 nt *Bubble* (2:1). XPA/DNA-Komplexpeak bei 1. b) RPA mit AAF geschädigtem 33mer und einer 13 nt *Bubble* (2:1). RPA/DNA-Komplexpeak bei 2. Verwendung der Gelfiltrationssäule SuperdexTM 200 Increase 10/300 GL (24 mL), Kristallisationspuffer II und 0 min Inkubationszeit. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. Die Hauptpeaks beider Chromatogramme haben ein Retentionsvolumina von 13.06 mL (1), und 10.06 mL (2). c) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion von 1 und SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 1. M: Proteinstandard PageRulerTM, 1: Eluat bei 13.06 mL. d) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion von 2 und SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 2. M: Proteinstandard PageRulerTM, 2: Eluat bei 10.04 mL.

a) XPA:RPA 1:1 Gelfiltration



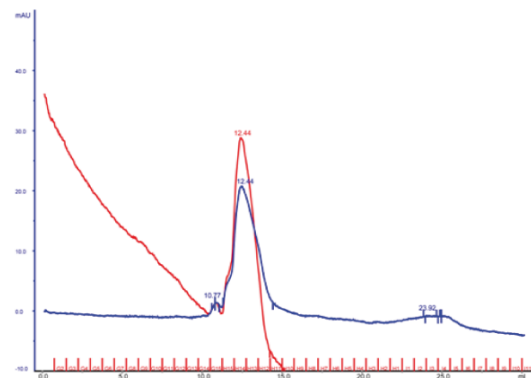
b) SDS-PAGE



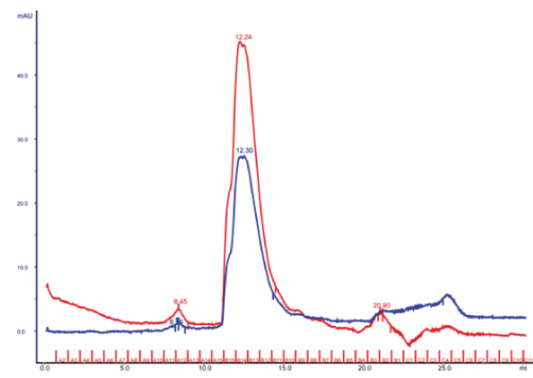
1

Abbildung 45: a) XPA mit RPA (1:1). XPA/RPA-Komplexpeak bei 1. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL), Kristallisationspuffer II und 0 min Inkubationszeit. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. Der Hauptpeak hat ein Retentionsvolumen von 11.84 mL (1). **b) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion von 1.** M: Proteinstandard PageRuler™, 1: Eluat bei 11.84 mL.

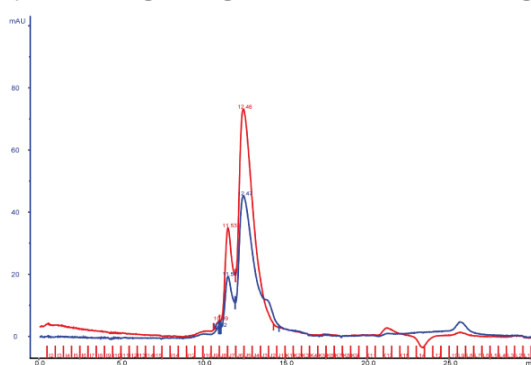
a) XPA mit AAF geschädigtes 25mer und 5'-20nt Überhang



b) XPA mit 25mer und 5'-20nt Überhang



c) XPA mit AAF geschädigtes 34mer und 5'-20nt Überhang



d) XPA mit 34mer und 5'-20nt Überhang

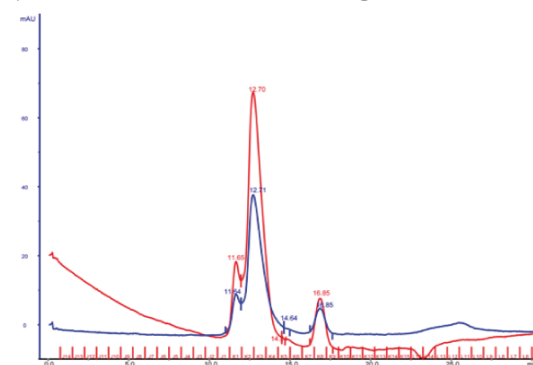
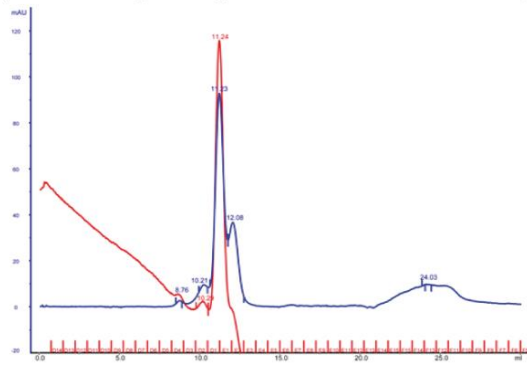
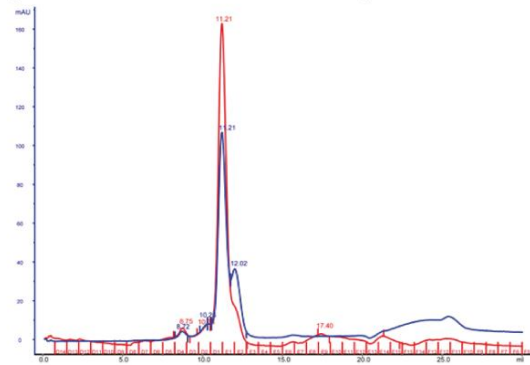


Abbildung 46: a) XPA mit AAF geschädigten 25mer und 20nt 5'-Überhang (2:1). XPA/DNA-Komplexpeak bei 12.44 mL. b) XPA mit einem 25mer und 20nt 5'-Überhang (2:1). XPA/DNA-Komplexpeak bei 12.24 mL. c) XPA mit AAF geschädigten 34mer und 20nt 5'-Überhang (2:1). XPA/DNA-Komplexpeak bei 12.00 mL. d) XPA mit einem 34mer und 20nt 5'-Überhang (2:1). XPA/DNA-Komplexpeak bei 12.70 mL. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL), Kristallisationspuffer II und 0 min Inkubationszeit. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. **SDS-PAGE und native Gele zu den einzelnen Chromatogrammen sind nicht mit dargestellt.**

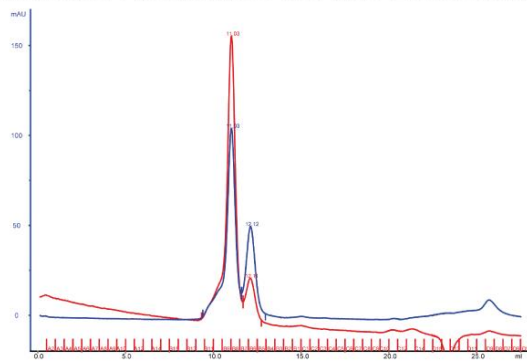
a) RPA mit AAF geschädigtes 25mer und 5'-20nt Überhang



b) RPA mit 25mer und 5'-20nt Überhang



c) RPA mit AAF geschädigtes 34mer und 5'-20nt Überhang



d) RPA mit 34mer und 5'-20nt Überhang

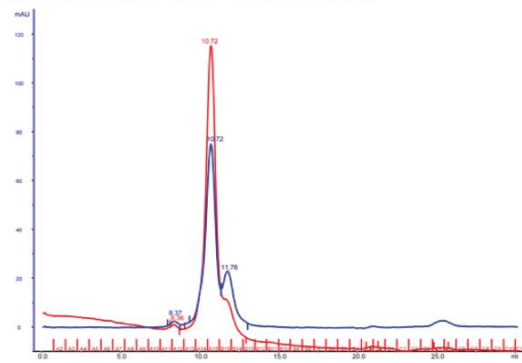


Abbildung 47: a) RPA mit AAF geschädigten 25mer und 20nt 5'-Überhang (2:1). RPA/DNA-Komplexpeak bei 11.24 mL. b) RPA mit einem 25mer und 20nt 5'-Überhang (2:1). RPA/DNA-Komplexpeak bei 11.21 mL. c) RPA mit AAF geschädigten 34mer und 20nt 5'-Überhang (2:1). RPA/DNA-Komplexpeak bei 11.03 mL. d) RPA mit einem 34mer und 20nt 5'-Überhang (2:1). RPA/DNA-Komplexpeak bei 10.72 mL. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL), Kristallisationspuffer II und 0 min Inkubationszeit. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. SDS-PAGE und native Gele zu den einzelnen Chromatogrammen sind nicht mit dargestellt.

6.5.3. Gelfiltrationsassay mit 15mer als Substrat

Gelfiltrationsassay von XPA mit RPA und verschiedenen 15mer-Substraten. Dabei zeigte sich, dass dieses DNA-Substrat sich nicht gut zur Komplexbildung eignet.

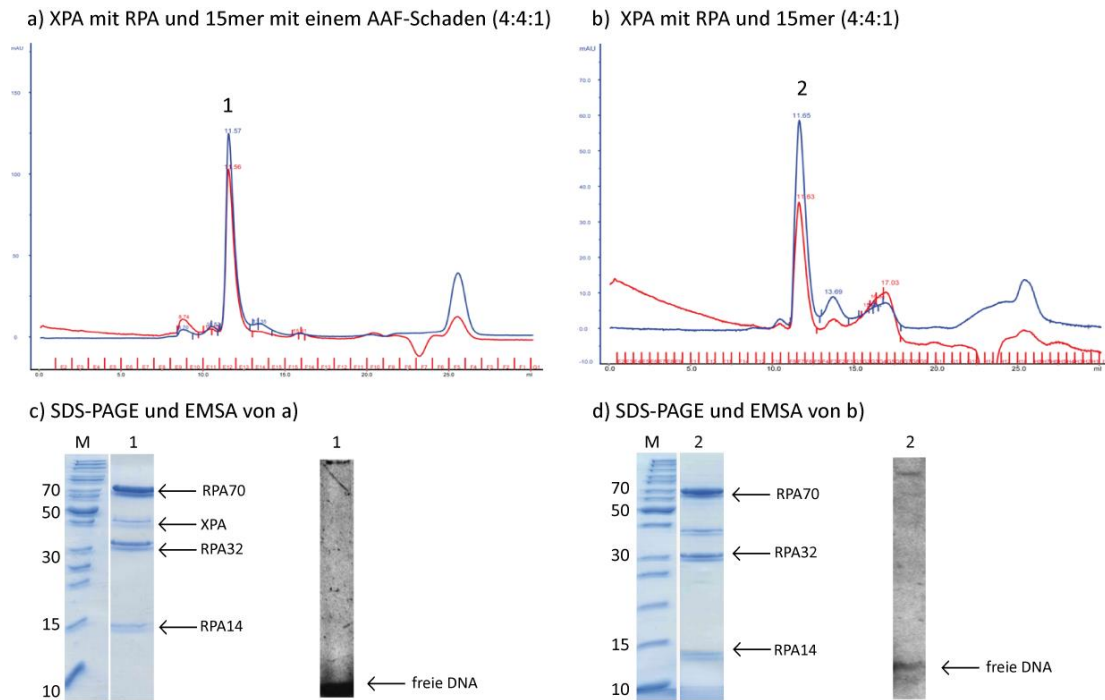


Abbildung 48: a) XPA mit RPA und 15mer mit einem AAF-Schaden (4:4:1). Ausbildung eines XPA/RPA-Komplexpeak bei 11.57 mL. Kaum DNA gebunden. b) XPA mit RPA und einem 15mer (4:4:1). Ausbildung eines XPA/RPA-Komplexpeak bei 11.65 mL. Kaum DNA gebunden. Verwendung der Gelfiltrationssäule SuperdexTM 200 Increase 10/300 GL (24 mL), Kristallisationspuffer II und 0 min Inkubationszeit. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. c) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion von 1 und SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 1. M: Proteinstandard PageRulerTM, 1: Eluat bei 11.57 mL. d) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion von 2 und SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 2. M: Proteinstandard PageRulerTM, 2: Eluat bei 11.65 mL. Nur ein Verhältnis gezeigt von den getesteten Verhältnissen 4:4:1, 2:2:1 und 1:1:1.

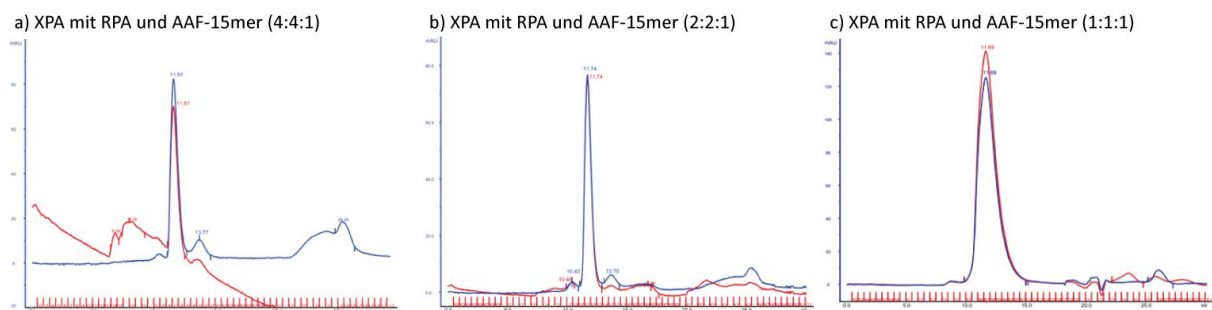


Abbildung 49: XPA mit RPA und 15mer mit einem AAF-Schaden im Verhältnis a) 4:4:1 b) 2:2:1 c) 1:1:1. Ausbildung eines XPA/RPA-Komplexes. Kaum DNA gebunden. Verwendung der Gelfiltrationssäule SuperdexTM 200 Increase 10/300 GL (24 mL), Kristallisationspuffer II und 0 min Inkubationszeit. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. SDS-PAGE und native Gele zu den einzelnen Chromatogrammen sind nicht dargestellt.

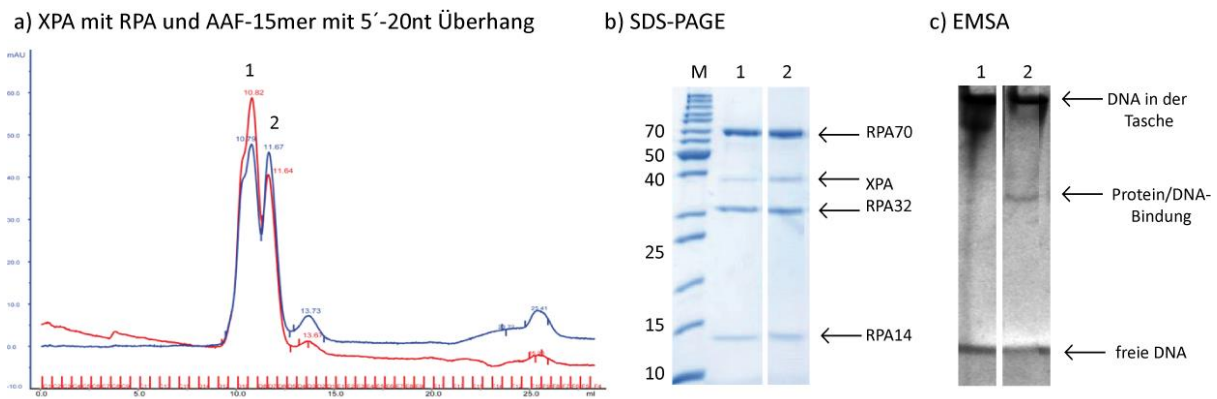


Abbildung 50: a) XPA mit RPA und AAF-geschädigtem 15mer mit 20nt 5'-Überhang (2:2:1). RPA/DNA-Komplexpeak bei 1. XPA/RPA-Komplexpeak bei 2 mit sehr wenig DNA. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL), Kristallisationspuffer II und 0 min Inkubationszeit. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. Die Hauptpeaks haben ein Retentionsvolumina von 10.82 mL (1) und 11.67 mL (2). **b)** Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion von 1 und 2. M: Proteinstandard PageRuler™, 1: Eluat bei 10.82 mL. 2: Eluat bei 11.67 mL. **c)** SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 1 und 2. 1: Eluat bei 10.82 mL. 2: Eluat bei 11.67 mL. Nur ein Verhältnis gezeigt von den getesteten Verhältnissen 2:2:1 und 1:1:1.

6.5.4. Gelfiltrationsassay mit 25mer als Substrat

XPA wurde mit RPA und verschiedenen 25mer-Substraten mittels Gelfiltration analysiert.

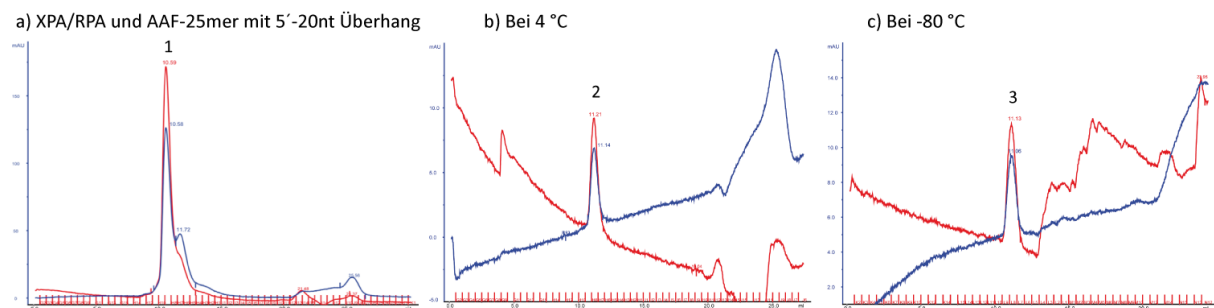
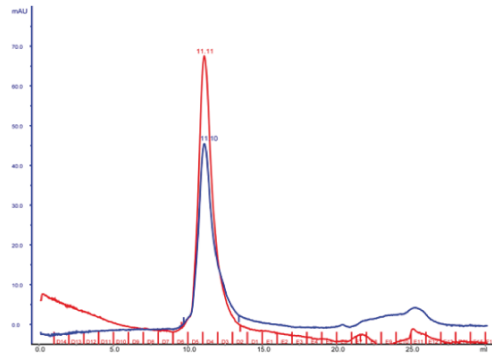
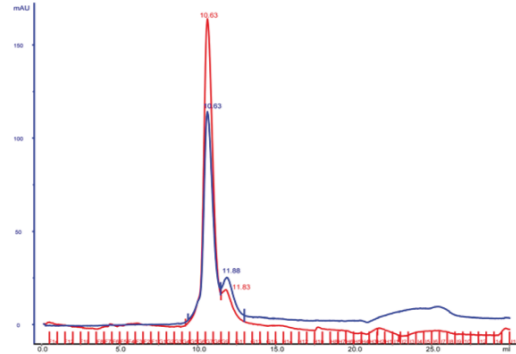


Abbildung 51: Stabilitätstests des XPA/RPA/DNA-Komplexes (2:2:1) mit einem AAF-geschädigten 25mer mit einem 20nt 5'-Überhang. **a)** Isolierung des Komplexes. **b)** Erneute Gelfiltration des isolierten Komplexes von Peak 1 aus a nach Lagerung bei 4 °C. **c)** Erneute Gelfiltration des isolierten Komplexes von Peak 1 aus a nach Lagerung bei -80 °C. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL) und Kristallisationspuffer II. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. Die Hauptpeaks beider Chromatogramme haben ein Retentionsvolumen von 11.52 mL (1) und 10.56 mL (2). SDS-PAGE und native Gele zu den einzelnen Chromatogrammen sind nicht mit dargestellt.

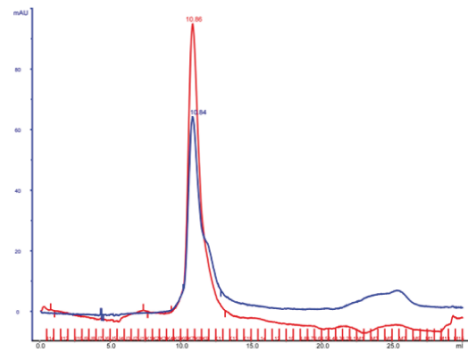
a) Gleichzeitig



b) RPA später



c) XPA später



d) DNA später

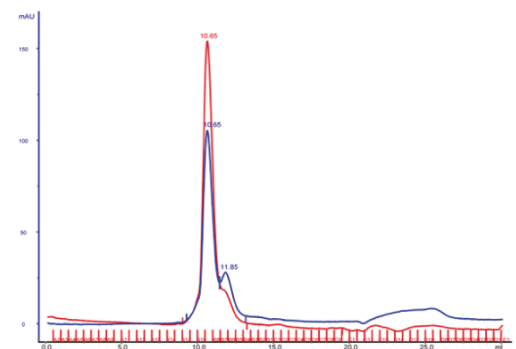
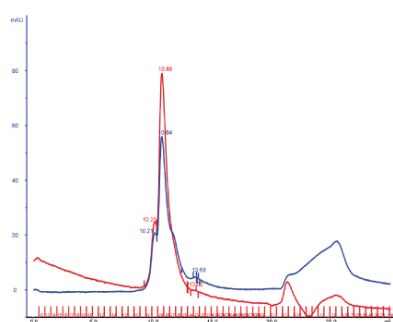


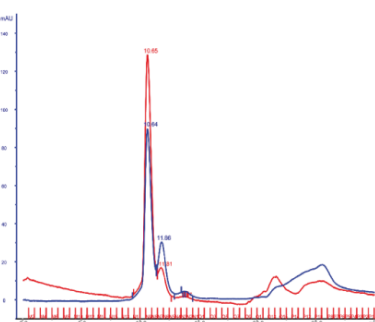
Abbildung 52: Vergleich der verschiedenen Zugabezeiten der Proteine und einem ungeschädigten 25mer mit einem 20nt 5'-Überhang zur Komplexbildung (Verhältnis 2:2:1 XPA:RPA:DNA). a) Gleichzeitig und 0 min Inkubation. b) Inkubation von XPA und DNA für 10 min; dann Zugabe von RPA. c) Inkubation von RPA und DNA für 10 min; dann Zugabe von XPA. d) Inkubation von XPA und RPA für 10 min; dann Zugabe von DNA. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL) und Kristallisationspuffer II. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. Kein Unterschied in der Komplexbildung. SDS-PAGE und native Gele zu den einzelnen Chromatogrammen sind nicht mit dargestellt.

XPA/RPA und AAF-25mer mit 5'20nt Überhang

a) 4:4:1



b) 2:2:1



c) 1:1:1

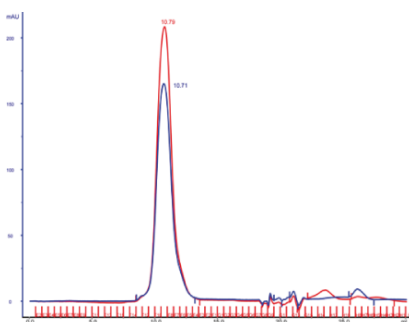
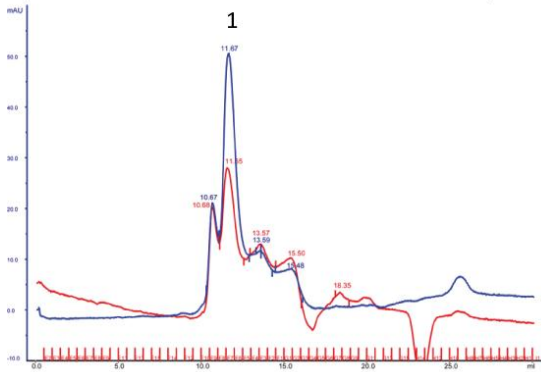
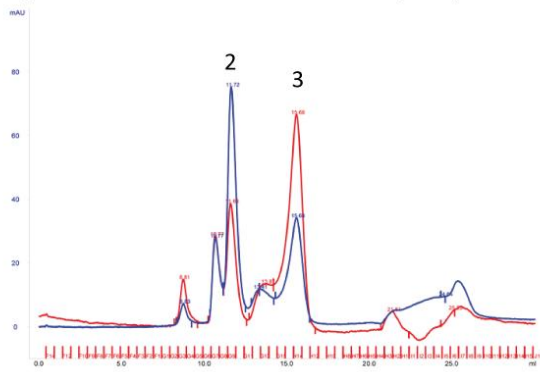


Abbildung 53: XPA mit RPA und 25mer mit einem AAF-Schaden im Verhältnis a) 4:4:1 b) 2:2:1 c) 1:1:1. Ausbildung eines XPA/RPADNA-Komplexes. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL), Kristallisationspuffer II und 0 min Inkubationszeit. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. SDS-PAGE und native Gele zu den einzelnen Chromatogrammen sind nicht mit dargestellt.

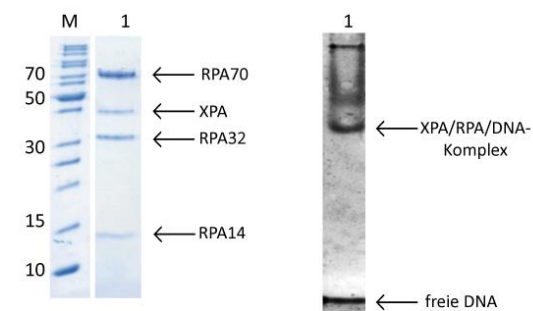
a) XPA/RPA und 25mer mit einem AAF-Schaden (2:2:1)



b) XPA/RPA und 25mer ohne Schaden (2:2:1)



c) SDS-PAGE und EMSA von a)



d) SDS-PAGE und EMSA von b)

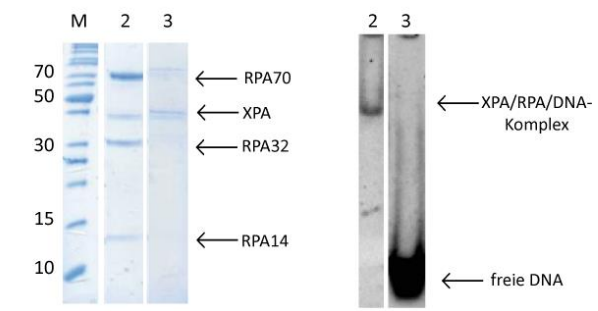
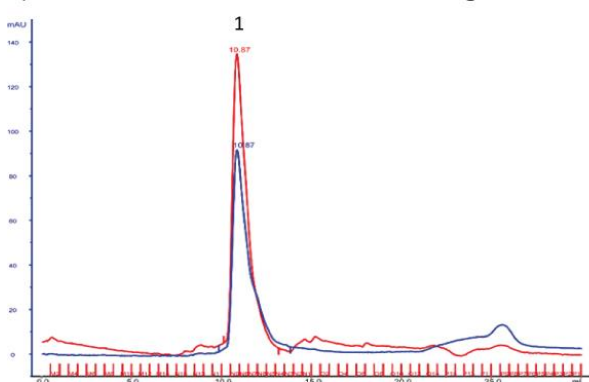
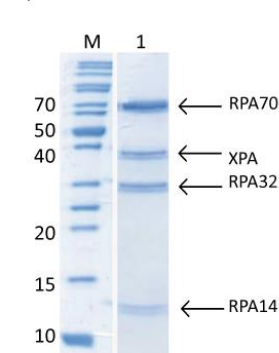


Abbildung 54: a) XPA mit RPA und 25mer mit einem AAF-Schaden (2:2:1). Ausbildung eines XPA/RPA-Komplexpeak bei 11.67 mL. Kaum DNA gebunden. b) XPA mit RPA und einem 25mer (2:2:1). Ausbildung eines XPA/RPA-Komplexpeak bei 11.72 mL. Kaum DNA gebunden. Verwendung der Gelfiltrationssäule SuperdexTM 200 Increase 10/300 GL (24 mL), Kristallisationspuffer II und 0 min Inkubationszeit. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. **c) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion und SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 1.** M: Proteinstandard PageRulerTM, 1: Eluat bei 11.67 mL. **d) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion und SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 2 und 3.** M: Proteinstandard PageRulerTM, 2: Eluat bei 11.72 mL, 3: Eluat bei 15.06 mL. **Nur ein Verhältnis gezeigt von den getesteten Verhältnissen 4:4:1 und 2:2:1.**

a) XPA/RPA und 25mer mit 5'-15nt Überhang



b) SDS-PAGE



c) EMSA

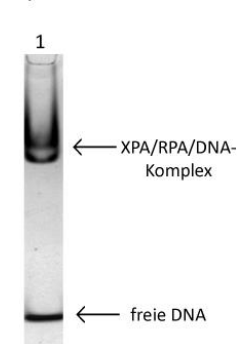
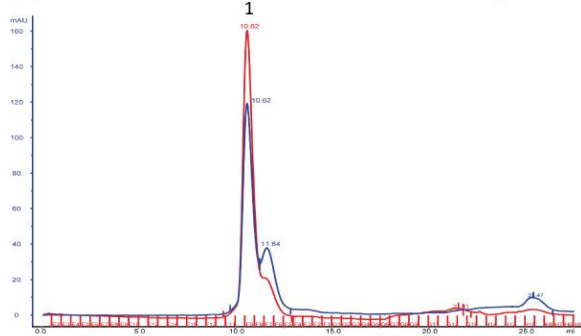
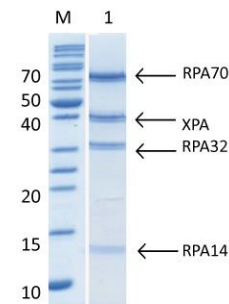


Abbildung 55: a) XPA mit RPA und 25mer mit 15nt 5'-Überhang (2:2:1). Komplexpeak bei 1. Verwendung der Gelfiltrationssäule SuperdexTM 200 Increase 10/300 GL (24 mL), Kristallisationspuffer II und 0 min Inkubationszeit. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. **b) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion von 1.** M: Proteinstandard PageRulerTM, 1: Eluat bei 10.67 mL. **c) SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 1.** 1: Eluat bei 10.67 mL. **Nur ein Verhältnis gezeigt von den getesteten Verhältnissen 4:4:1 und 2:2:1.**

a) XPA/RPA und AAF-25mer mit 5'-10nt Überhang



b) SDS-PAGE



c) EMSA

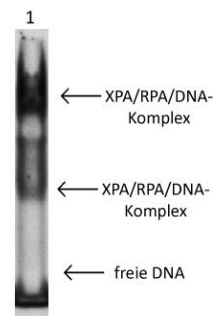
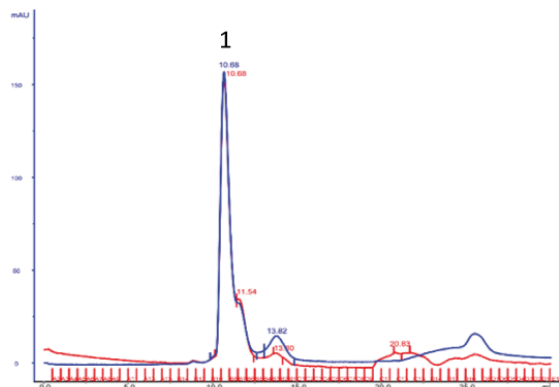


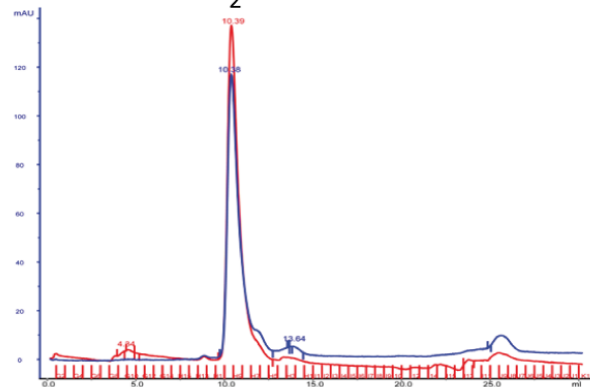
Abbildung 56: a) XPA mit RPA und 25mer mit 10nt 5'-Überhang (2:2:1). Komplexpeak bei 1. Verwendung der Gelfiltrationssäule SuperdexTM 200 Increase 10/300 GL (24 mL), Kristallisationspuffer II und 0 min Inkubationszeit. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. b) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion von 1. M: Proteinstandard PageRulerTM, 1: Eluat bei 10.62 mL. c) SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 1. Verschiedene Spezies im nativen Gel. 1: Eluat bei 10.62 mL. Nur ein Verhältnis gezeigt von den getesteten Verhältnissen 4:4:1 und 2:2:1.

XPA/RPA und 25mer mit 3 Mismatches an beiden Enden (2:2:1)

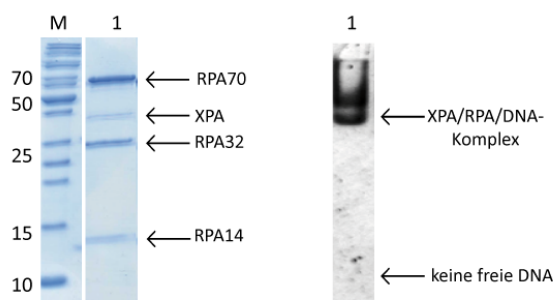
a) mit AAF-Schaden



b) ohne DNA-Schaden



c) SDS-PAGE und EMSA von a)



d) SDS-PAGE und EMSA von b)

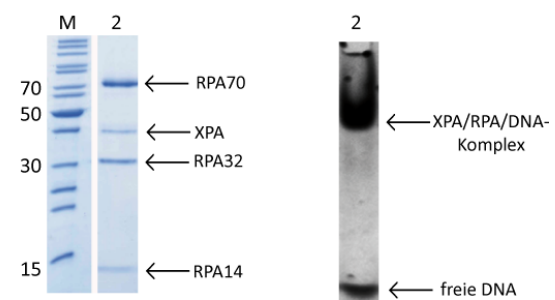


Abbildung 57: XPA/RPA mit einem 25mer mit 3 ungepaarten Fehlstellen an beiden Enden (2:2:1); a) mit einem AAF-Schaden (2:2:1). Ausbildung eines XPA/RPA/DNA-Komplexpeak bei 11.68 mL. Kaum XPA gebunden. b) ohne DNA Schaden. Ausbildung eines XPA/RPA/DNA-Komplexpeak bei 10.39 mL. Kaum XPA gebunden. Verwendung der Gelfiltrationssäule SuperdexTM 200 Increase 10/300 GL (24 mL), Kristallisationspuffer II und 0 min Inkubationszeit. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. c) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion und SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 1. M: Proteinstandard PageRulerTM, 1: Eluat bei 11.68 mL. d) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion und SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 2. M: Proteinstandard PageRulerTM, 2: Eluat bei 11.39 mL. Nur ein Verhältnis gezeigt von den getesteten Verhältnissen 4:4:1, 2:2:1 und 1:1:1.

6.5.5. Gelfiltrationsassay mit 34mer als Substrat

XPA wurde mit RPA und verschiedenen 34mer-Substraten mittels Gelfiltration analysiert.

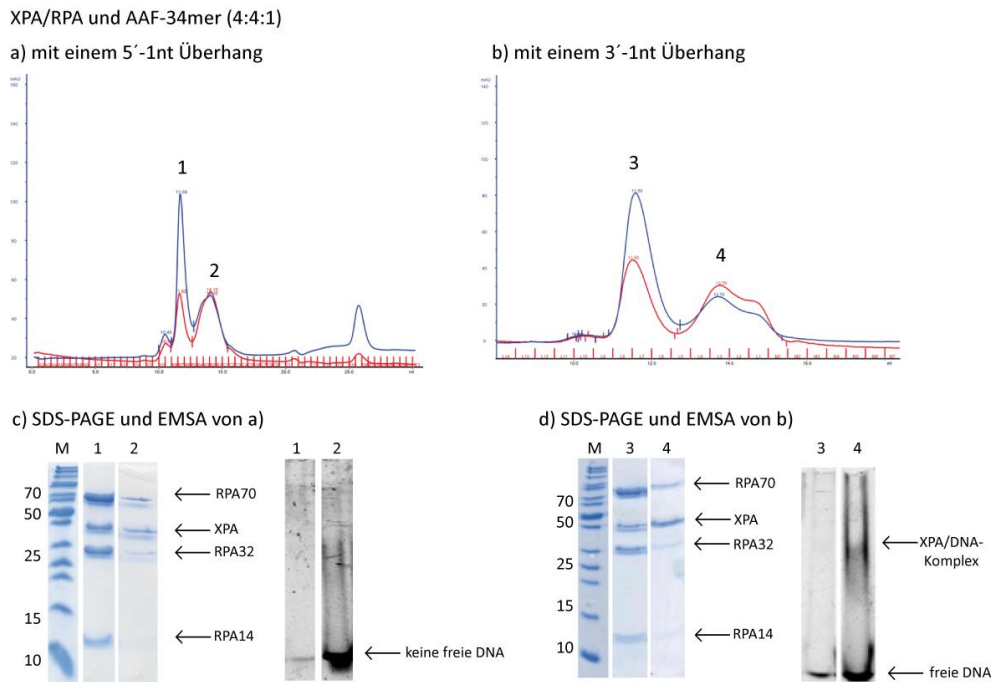


Abbildung 58: XPA/RPA mit einem 34mer mit einem AAF-Schaden (4:4:1); a) mit einem 1nt 5'-Überhang. Ausbildung eines XPA/RPA-Komplexpeak bei 1. DNA-Peak bei 2. b) mit einem 1nt 3'-Überhang. Ausbildung eines XPA/RPA-Komplexpeak bei 3. XPA/DNA-Peak bei 4. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL), Kristallisationspuffer II und 0 min Inkubationszeit. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. c) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion und SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 1 und 2. M: Proteinstandard PageRuler™, 1: Eluat bei 11.68 mL, 2: Eluat bei 14.10 mL. d) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion und SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 3 und 4. M: Proteinstandard PageRuler™, 3: Eluat bei 11.59 mL, 4: Eluat bei 13.76 mL.

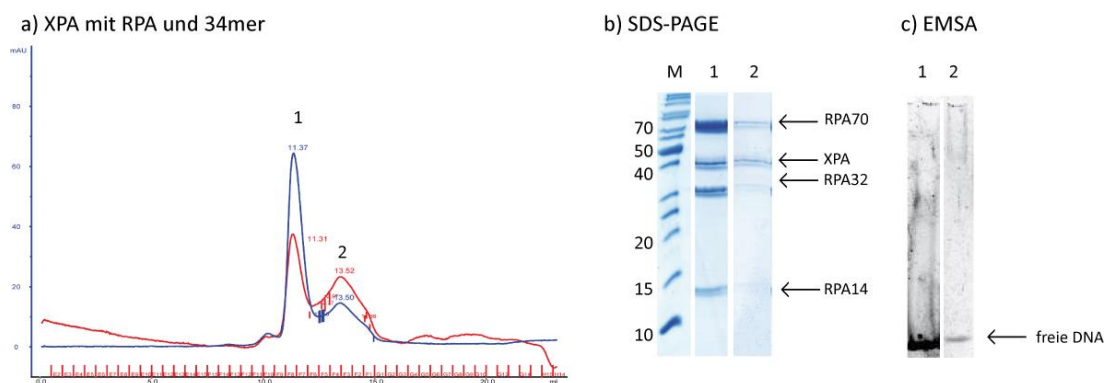


Abbildung 59: a) XPA mit RPA und einem ungeschädigten 34mer (4:4:1). XPA/RPA-Peak bei 1. Kaum DNA gebunden. DNA-Peak bei 2. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL). x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]. blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. b) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion von 1 und 2. M: Proteinstandard PageRuler™, 1: Eluat bei 11.37 mL, 2: Eluat bei 11.52 mL. c) SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 1 und 2. 1: Eluat bei 11.37 mL, 2: Eluat bei 11.52 mL.

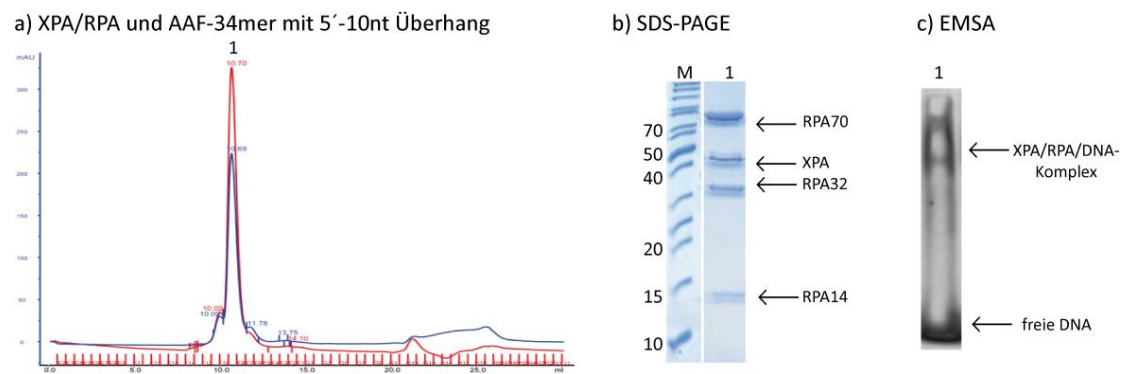


Abbildung 60: a) XPA mit RPA und einem AAF-geschädigten 34mer mit einem 10nt 5'-Überhang (2:2:1). XPA/RPA/DNA-Peak bei 1. Kaum DNA gebunden. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL). x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. b) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion von 1. M: Proteinstandard PageRuler™, 1: Eluat bei 11.70 mL. c) SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 1. Verschiedene Spezies im nativen Gel. 1: Eluat bei 11.70 mL. Nur ein Verhältnis gezeigt von den getesteten Verhältnissen 4:4:1 und 2:2:1.

6.5.6. Gelfiltrationsassay mit *Bubble*-DNA als Substrat

XPA wurde mit RPA und verschiedenen *Bubble*-Substraten mittels Gelfiltration analysiert.

XPA/RPA und 34mer mit 13nt *Bubble* (2:2:1)

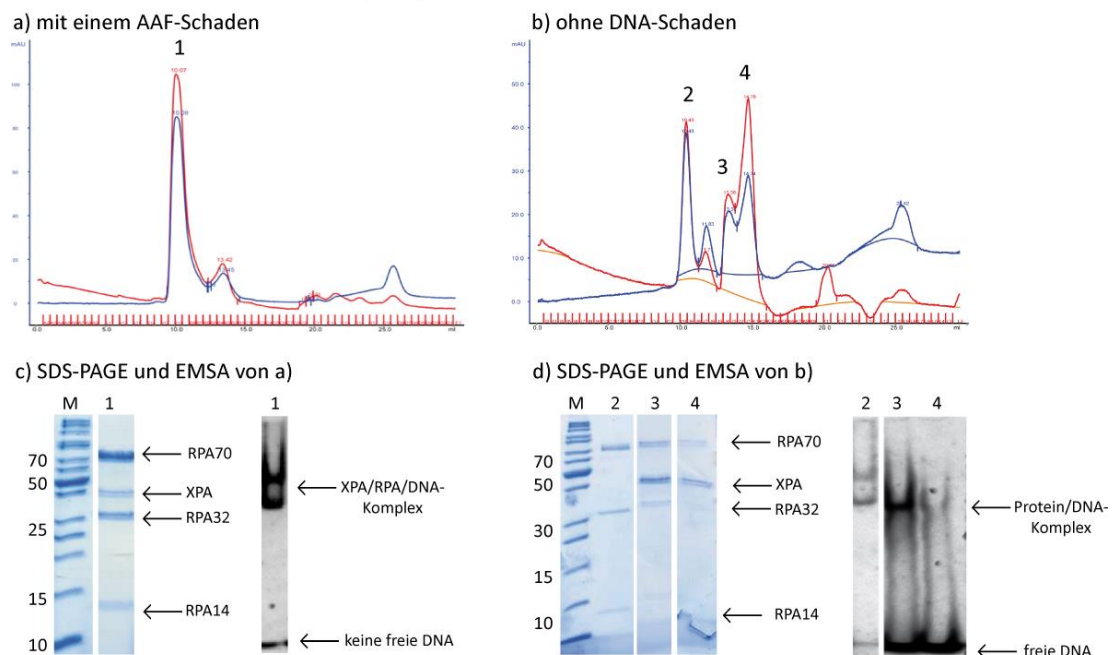


Abbildung 61: XPA/RPA mit einem 34mer mit einer 13nt *Bubble*, welche durch 13 ungepaarte Fehlstellen entsteht (2:2:1); a) mit einem AAF-Schaden (2:2:1). Ausbildung eines XPA/RPA/DNA-Komplexpeaks bei 1. b) ohne DNA Schaden. Ausbildung eines RPA/DNA-Komplexpeaks bei 2, XPA/DNA-Komplexes bei 3 und DNA-Peak bei 4. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL), Kristallisationspuffer II und 0 min Inkubationszeit. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. c) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion und SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 1. M: Proteinstandard PageRuler™, 1: Eluat bei 10.07 mL. d) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion und SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 2, 3 und 4. M: Proteinstandard PageRuler™, 2: Eluat bei 10.43 mL, 3: Eluat bei 13.36 mL, 4: Eluat bei 14.76 mL. Nur ein Verhältnis gezeigt von den getesteten Verhältnissen 4:4:1, 2:2:1 und 1:1:1.

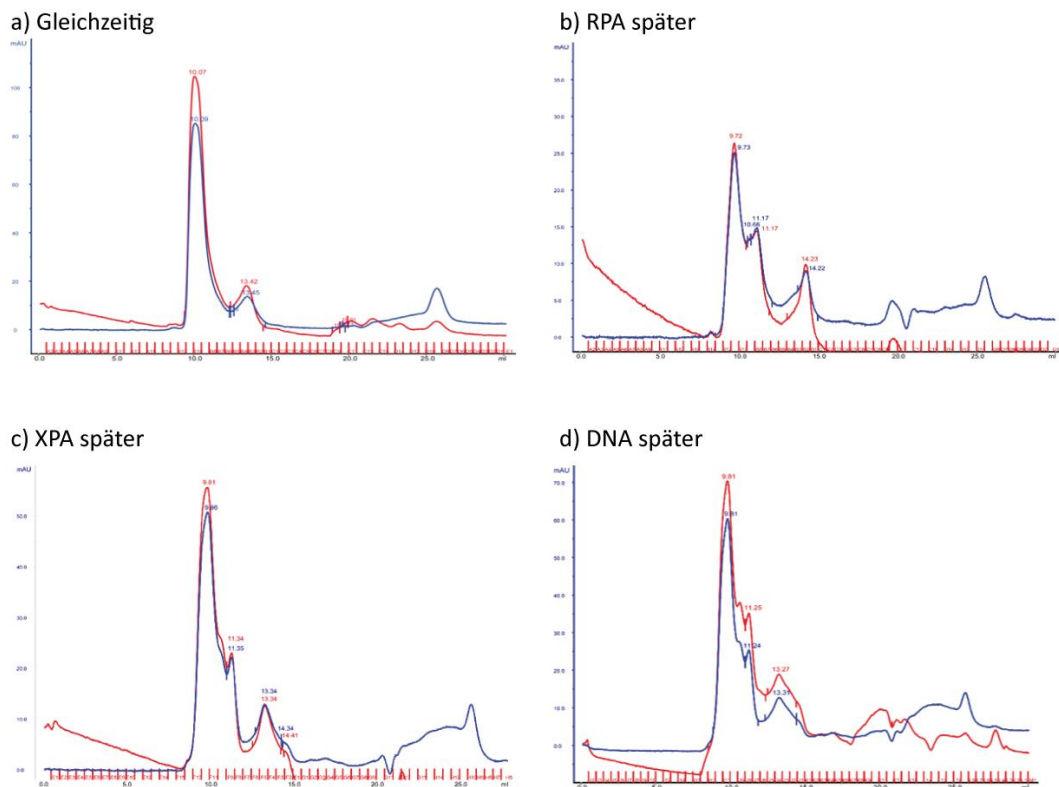


Abbildung 62: Vergleich der verschiedenen Zugabezeiten der Proteine und einem AAF-geschädigten 34mer mit einer 13nt *Bubble* (2:2:1). a) Gleichzeitig und 0 min Inkubation. b) Inkubation von XPA und DNA für 10 min; dann Zugabe von RPA. c) Inkubation von RPA und DNA für 10 min; dann Zugabe von XPA. d) Inkubation von XPA und RPA für 10 min; dann Zugabe von DNA. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL) und Kristallisationspuffer II. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. **Leichter Unterschied in der Komplexbildung.** SDS-PAGE und native Gele zu den einzelnen Chromatogrammen sind nicht mit dargestellt.

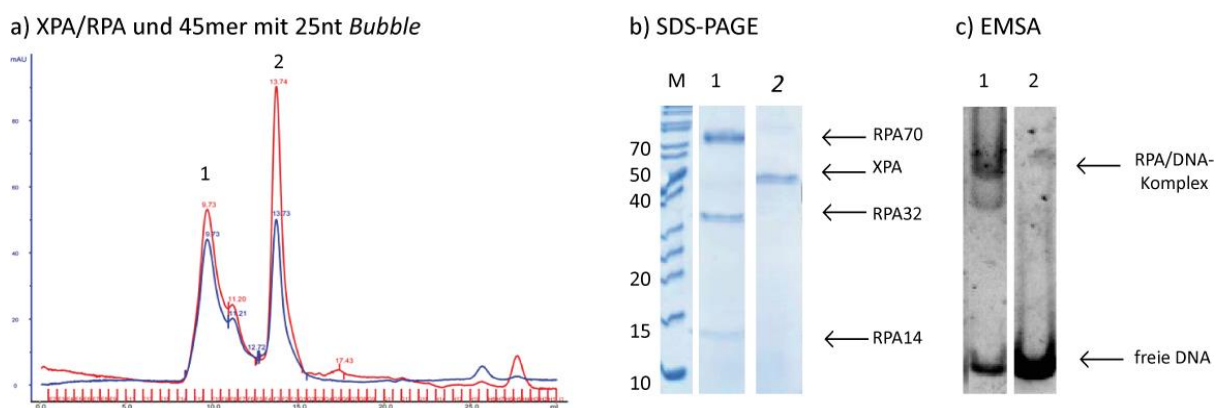


Abbildung 63: a) XPA mit RPA und ungeschädigtem 45mer mit einer 25nt *Bubble* (2:2:1). RPA/DNA-Komplexpeak bei 1. XPA/DNA-Komplexpeak bei 2, aber nicht stabil. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL), Kristallisationspuffer II und 0 min Inkubationszeit. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. **b) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion von 1 und 2.** M: Proteinstandard PageRuler™, 1: Eluat bei 9.73 mL, 2: Eluat bei 13.74 mL. **c) SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 1.** Verschiedene Spezies im nativen Gel. 1: Eluat bei 9.73 mL, 2: Eluat bei 13.74 mL. **Nur ein Verhältnis gezeigt von den getesteten Verhältnissen 4:4:1, 2:2:1 und 1:1:1.**

a) mit einem AAF-Schaden

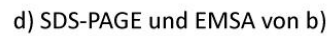


Abbildung 64: XPA/RPA mit einem 39mer mit einer 19nt *Bubble*, welche durch 19 ungepaarte Fehlstellen entsteht (2:2:1); a) mit einem AAF-Schaden (2:2:1). Ausbildung eines XPA/RPA/DNA-Komplexpeaks bei 1. **b) ohne DNA Schaden.** Ausbildung eines RPA/DNA-Komplexpeak bei 2, XPA/DNA-Komplexes bei 3 und DNA-Peak bei 4. Verwendung der Gelfiltrationssäule Superdex™ 200 Increase 10/300 GL (24 mL), Kristallisationspuffer II und 0 min Inkubationszeit. x-Achse: Retentionsvolumen [mL]; y-Achse: Absorption [mAU]; blau: Absorption bei 280 nm, rot: Absorption bei 260 nm. **c) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion und SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 1.** M: Proteinstandard PageRuler™, 1: Eluat bei 111.11 mL. **d) Coomassie gefärbtes 15%iges SDS-PAGE der Protein enthaltende Fraktion und SYBR Green I gefärbtes natives Gel der DNA enthaltende Fraktion von 2, 3 und 4.** M: Proteinstandard PageRuler™, 2: Eluat bei 9.94 mL, 3: Eluat bei 11.36 mL, 4: Eluat bei 14.33 mL. Nur ein Verhältnis gezeigt von den getesteten Verhältnissen 4:4:1, 2:2:1 und 1:1:1.

6.6. EMSA-Bindungsstudien

Zu den Gelfiltrationsstudien wurden zeitgleich EMSA-Bindungsstudien durchgeführt, um zu testen, ob die einzelnen Proteine an die DNA binden oder ob der Komplex im EMSA-Gel erkennbar ist.

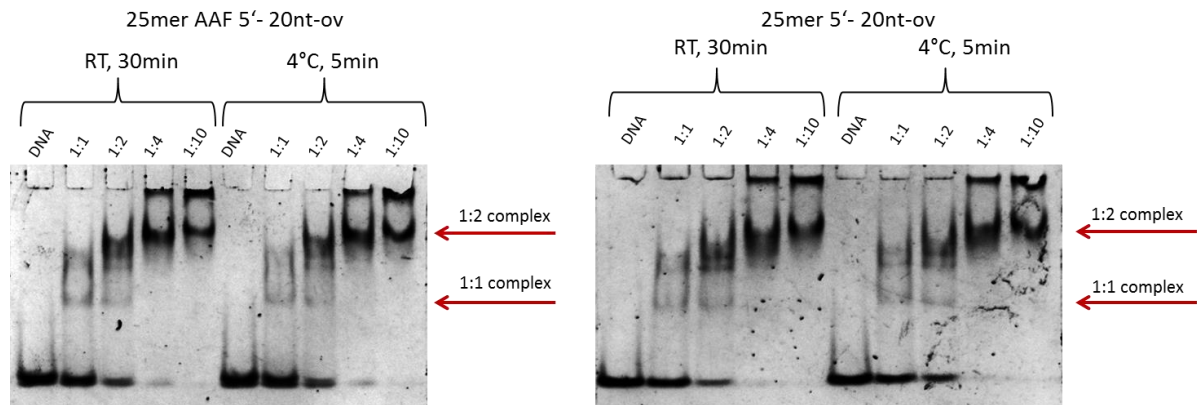


Abbildung 65: Kleine EMSA-Gele von XPA mit 25mer AAF 20nt 5'-Überhang und 25mer 20nt 5'-Überhang DNA. Es wurden unterschiedliche Inkubationszeiten (5 min und 30 min) und Temperaturen (RT und 4 °C) getestet. XPA wurde mit steigender Konzentration zu 10 pmol DNA hinzugegeben. Laufbedingungen: 45 min, 100 V und 4 °C.

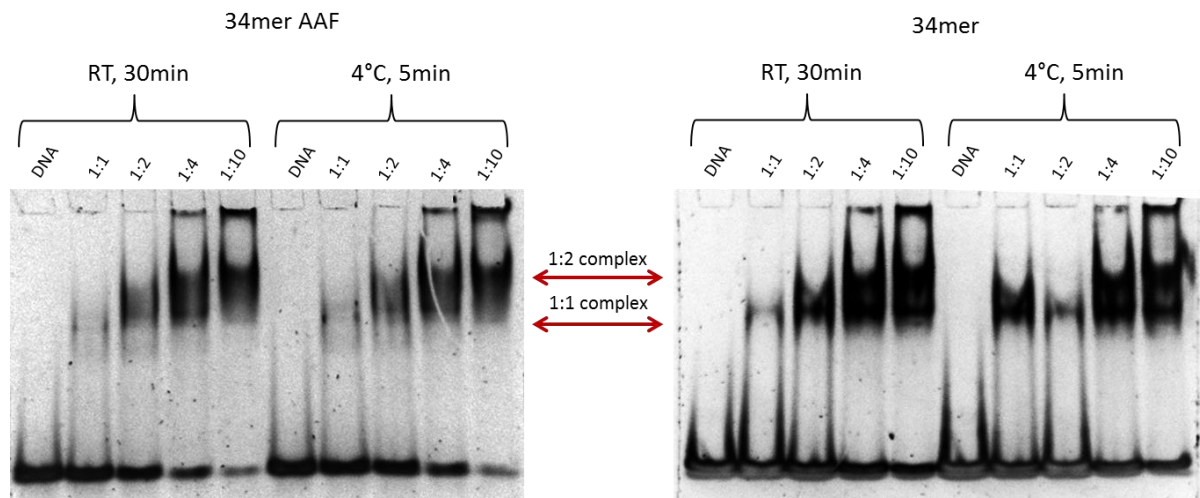


Abbildung 66: Kleine EMSA-Gele von XPA mit einem AAF-geschädigten 34mer und einem 34mer. Es wurden unterschiedliche Inkubationszeiten (5 min und 30 min) und Temperaturen (RT und 4 °C) getestet. XPA wurde mit steigender Konzentration zu 10 pmol DNA hinzugegeben. Laufbedingungen: 45 min, 100 V und 4 °C.

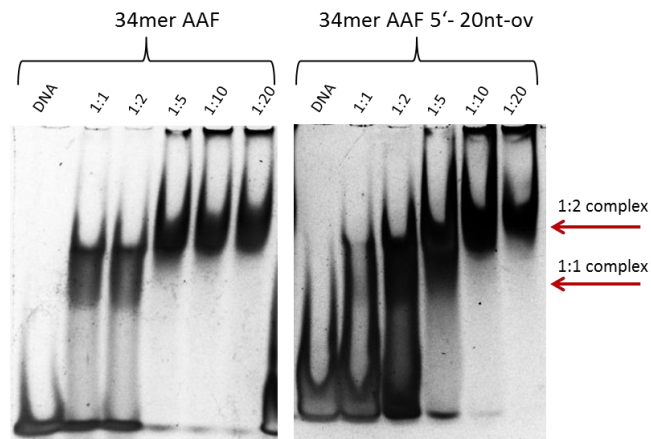


Abbildung 67: Kleine EMSA-Gele von XPA mit einem AAF-geschädigten 34mer und einem AAF-geschädigten 34mer mit einem 20nt 5'-Überhang. XPA wurde mit steigender Konzentration zu 10 pmol DNA hinzugegeben und 30 min bei RT inkubiert. Laufbedingungen: 45 min, 100 V und 4 °C.

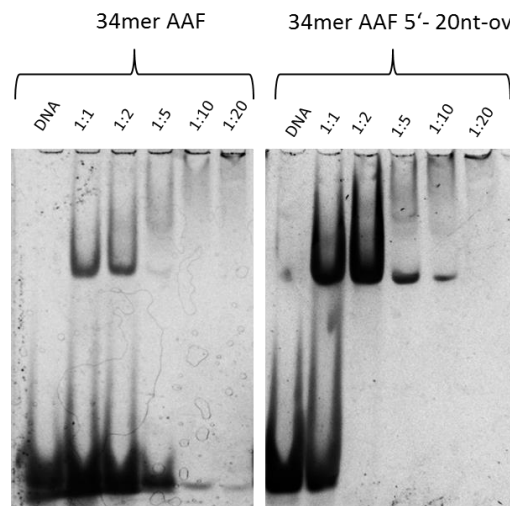


Abbildung 68: Kleine EMSA-Gele von RPA mit einem AAF-geschädigten 34mer und einem AAF-geschädigten 34mer mit einem 20nt 5'-Überhang. RPA wurde mit steigender Konzentration zu 10 pmol DNA hinzugegeben und 30 min bei RT inkubiert. Laufbedingungen: 45 min, 100 V und 4 °C.

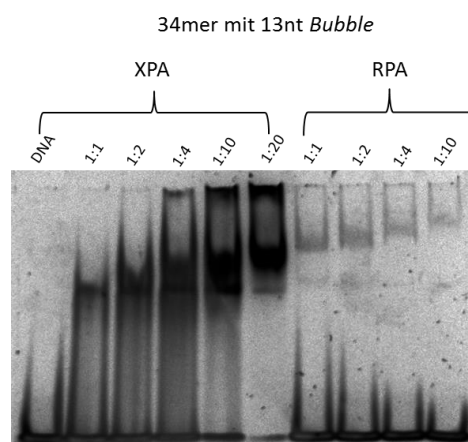


Abbildung 69: Kleines EMSA-Gel von XPA und RPA mit einem 34mer mit einer 13nt *Bubble*. Die Proteine wurde jeweils mit steigender Konzentration zu 10 pmol DNA hinzugegeben und 60 min bei 4 °C inkubiert. Laufbedingungen: 45 min, 100 V und 4 °C.

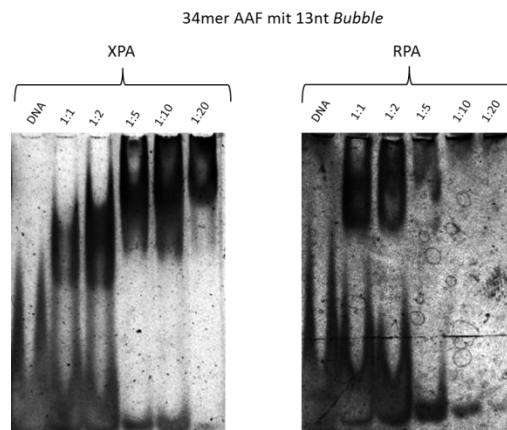


Abbildung 70: Kleine EMSA-Gele von XPA und RPA mit einem AAF-geschädigten 34mer mit einer 13nt *Bubble*. Die Proteine wurde jeweils mit steigender Konzentration zu 10 pmol DNA hinzugegeben und 60 min bei 4 °C inkubiert. Laufbedingungen: 45 min, 100 V und 4 °C.

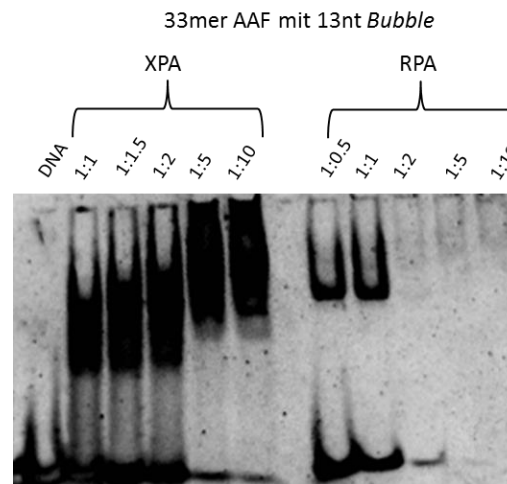


Abbildung 71: Kleines EMSA-Gel von XPA und RPA mit einem AAF-geschädigten 33mer mit einer 13nt *Bubble*. Die Proteine wurde jeweils mit steigender Konzentration zu 10 pmol DNA hinzugegeben und 60 min bei 4 °C inkubiert. Laufbedingungen: 45 min, 100 V und 4 °C.

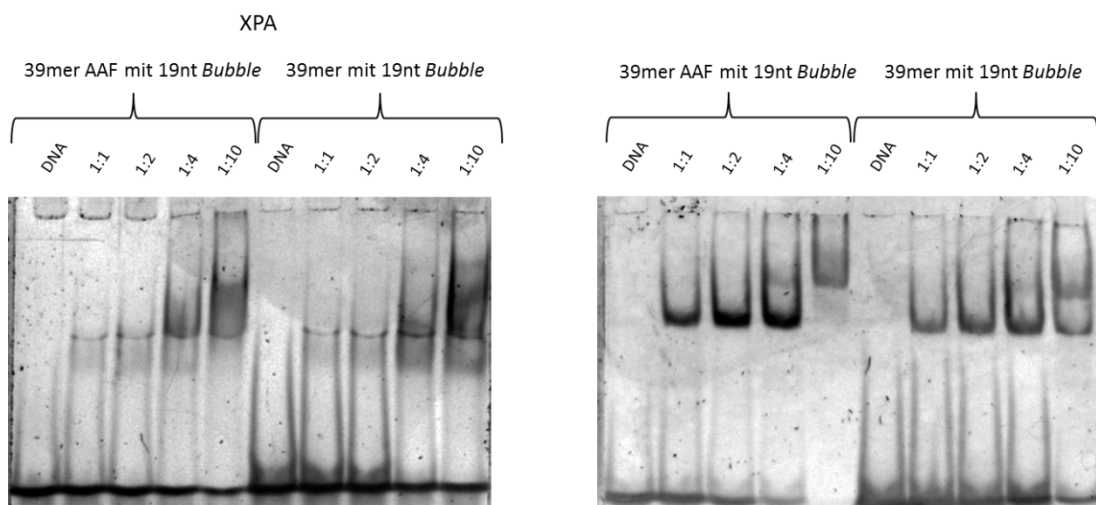


Abbildung 72: Kleine EMSA-Gele von XPA und RPA mit einem ungeschädigten und AAF-geschädigten 39mer mit einer 19nt *Bubble*. Die Proteine wurde jeweils mit steigender Konzentration zu 10 pmol DNA hinzugegeben und 60 min bei 4 °C inkubiert. Laufbedingungen: 45 min, 100 V und 4 °C.

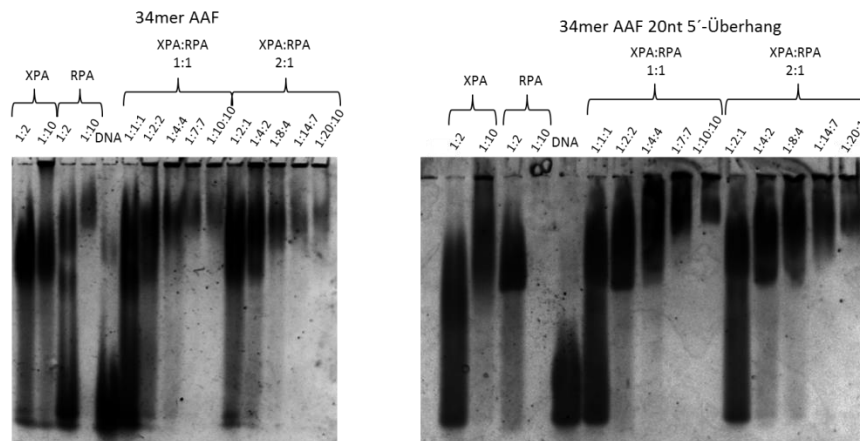


Abbildung 73: Große EMSA-Gele von XPA und RPA einzeln und im Komplex mit einem AAF-geschädigten 34mer und einem AAF-geschädigten 34mer mit einem 20nt 5'-Überhang. Die Proteine wurde jeweils mit steigender Konzentration zu 30 pmol DNA hinzugegeben und 30 min bei RT inkubiert. Laufbedingungen: 3.5 h und 10 mA.

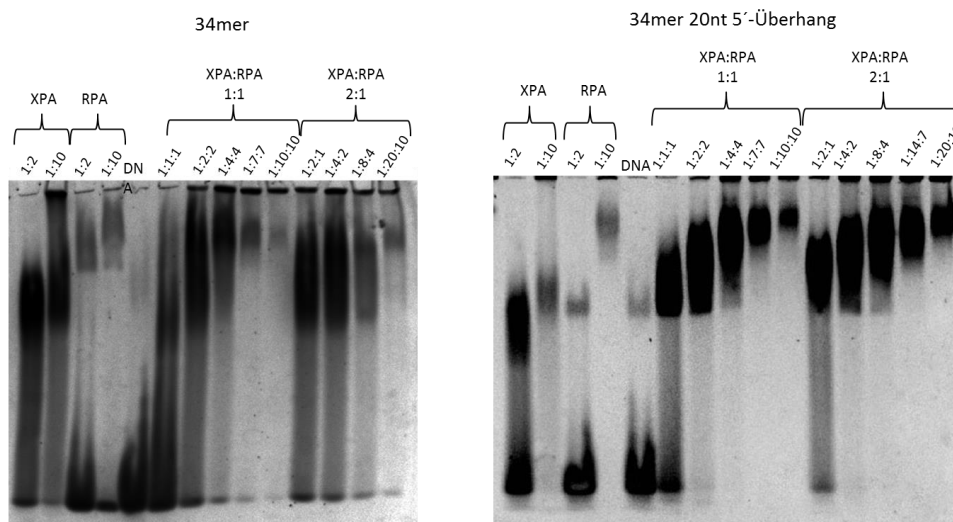


Abbildung 74: Große EMSA-Gele von XPA und RPA einzeln und im Komplex mit einem 34mer und einem 34mer mit einem 20nt 5'-Überhang. Die Proteine wurde jeweils mit steigender Konzentration zu 30 pmol DNA hinzugegeben und 30 min bei RT inkubiert. Laufbedingungen: 3.5 h und 10 mA.

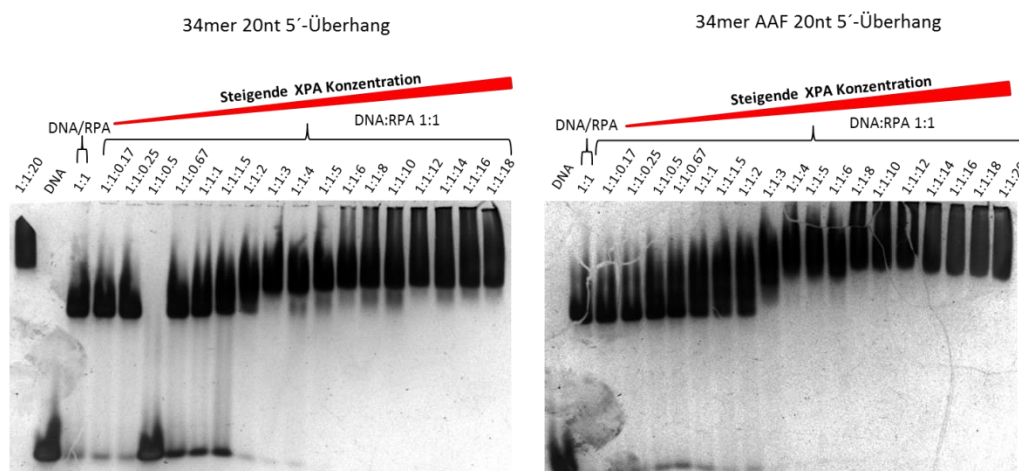


Abbildung 75: Titrationsgele: Große EMSA-Gele von XPA und RPA im Komplex mit einem 34mer 20nt 5'-Überhang und einem AAF-geschädigten 34mer 20nt 5'-Überhang. XPA wurde jeweils mit steigender Konzentration zu 20 pmol DNA und RPA hinzugegeben und 1 h bei 4 °C inkubiert. Laufbedingungen: 7 h, 10 mA und 7 °C.

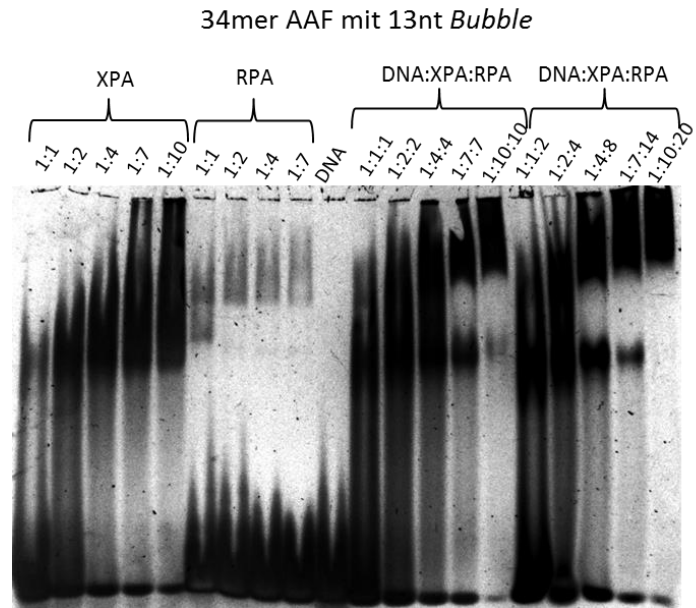


Abbildung 76: Großes EMSA-Gel von XPA und RPA einzeln und im Komplex mit einem 34mer mit einer 13nt *Bubble*. Die Proteine wurde jeweils mit steigender Konzentration zu 20 pmol DNA hinzugegeben und 30 min bei RT inkubiert. Laufbedingungen: 3.5 h, 7 °C und 10 mA.

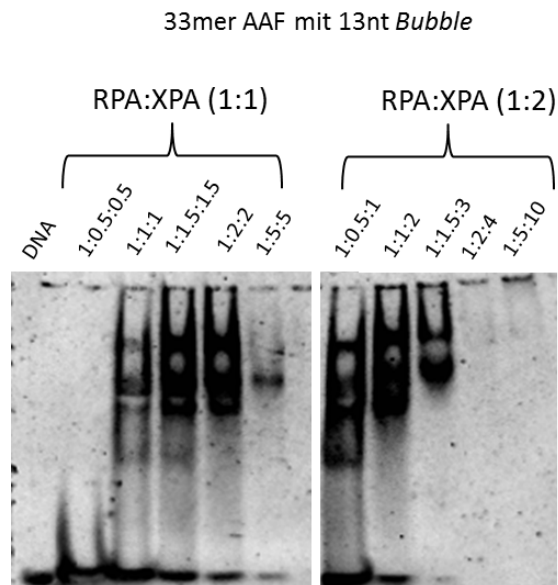


Abbildung 77: Kleines EMSA-Gel von XPA und RPA im Komplex mit einem AAF-geschädigten 33mer 13nt *Bubble*. Die Proteine wurde jeweils mit steigender Konzentration zu 10 pmol DNA hinzugegeben und 60 min bei 4 °C inkubiert. Laufbedingungen: 45 min, 100 V und 4 °C.

6.7. Daten der Kristallstrukturen

Tabelle 25: Experimentelle Daten der Ko-Kristallstruktur von Rad14 mit einem AAF-geschädigten 15mer mit einer 3nt Bubble. Statistiken mit der höchsten Auflösung werden in Klammern dargestellt. Werte, die mit einem * markiert wurden, wurden mit dem Molprobity Webserver (<http://molprobity.biochem.duke.edu>) kalkuliert.

	Rad14-AAF-dG mit 3nt <i>Bubble</i>
Wavelength (Å)	0.966
Resolution range	38.1 - 2.5 (2.6 - 2.5)
Space group	P4 ₁
Unit cell	53.8 53.8 131.7
Total reflections	47,023 (4,316)
Unique reflections	12,766 (1,250)
Multiplicity	3.7 (3.5)
Completeness (%)	0.98 (0.98)
Mean I/sigma(I)	10.7 (0.7)
Wilson B-factor	76.6
R-merge	0.07 (1.21)
R-meas	0.08 (1.42)
CC1/2	0.996 (0.41)
CC*	0.999 (0.76)
Reflections used in refinement	12,708 (1,208)
Reflections used for R-free	636 (60)
R-work	22.5
R-free	25.7
CC(work)	0.87 (0.58)
CC(free)	0.77 (0.68)
Number of non-hydrogen atoms	
Protein	1,868
DNA	1,100
Zinc-ions	2

Waters	3
Protein residues	227
RMSZ (bonds)	0.014
Bad bonds (protein)*	0
Bad bonds (DNA)*	1.63
RMSZ (angles)	0.17
Bad angles (protein)*	0.08
Bad angles (DNA)*	0.58
Ramachandran favored (%)*	98.2
Ramachandran allowed (%)*	1.8
Ramachandran outliers (%)*	-
Rotamer outliers (%)*	3
Clashscore*	3.2
Average B-factor	
Protein	79.5
DNA	121.8
Zinc ions	63.1
Waters	
Number of TLS groups	20

7. Abkürzungsverzeichnis

3'-Ende	3-Strich der Oligonukleotide
5'-Ende	5-Strich der Oligonukleotide
äq.	Äquivalente
AAF	Acetylaminofluoren
AF	Aminofluoren
Amp	Ampicillin
AP	Auftragspuffer
APS	Ammoniumperoxodisulfat
Asp	Asparaginsäure
ATR	<i>ATM-Rad3-related</i>
AU	Absorption units
B[a]P	Benzo[a]pyren
Bp	Basenpaare
BER	Basenexzisionsreparatur
<i>Bulky</i>	sperrig
c	Konzentration
cDNA	komplementäre DNA
c _{END}	Endkonzentration
ca.	Circa
Carb	Carbenicillin
Cam	Chloramphenicol
CH ₃ CN	Acetonitril
CPG	<i>Controlled pore glass</i>
CS	<i>Cockayne</i> Syndrom
CSA	<i>Cockayne</i> Syndrom A
CSB	<i>Cockayne</i> Syndrom B
CV	Säulenvolumen, <i>column volumn</i>
Da	Dalton
DBD	DNA-bindende-Domäne
dC	2'-Desoxycytosin
DCA	Dichloressigsäure
dd	Bidestilliert
dG	2'-Desoxyguanosin

C8-AAF-dG	N-(Desoxyguanosin-8-yl)-2-acetylaminofluoren
DNA	Desoyribonukleinsäure (<i>deoxyribonucleic acid</i>)
ds	Doppelsträngig
dsDNA	doppelsträngige DNA
DSD	Doppelstrangbruch
dT	2'-Desoxythymidin
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylendiamintetraacetat
EGFP	<i>Enhanced green fluorescent protein</i>
EM	Elektronenmikroskopie
EMSA	<i>Electromobility shift assay</i>
<i>et al.</i>	<i>et alii</i>
EtOH	Ethanol
FACS	<i>Fluorescence-activated cell sorting</i>
FITC	Fluoresceinisothiocyanat
FPLC	<i>fast protein liquid chromatography</i>
g	Gramm
GEBlue	Cibacron Blue 3G Farbstoff grkoppelt an einer Sepharose Matrix
GGR	Globale Genomreparatur (<i>global genome repair</i>)
H ₂ O	Wasser
HAP	Hydroxylapatit
HEPES	2-(4-(2-Hydroxyethyl)-1-piperazinyl)-ethansulfonsäure
His	Histidin
HPA	3-Hydroxypicolin-Säure
HPLC	Hochleistungsflüssigkeitschromatographie
HR	Homologe Rekombination
<i>i</i>	<i>iso</i>
IPTG	Isopropyl- β -D-thiogalactopyranosid
λ	Wellenlänge
K	Lysin
Kan	Kanamycin
KCl	Kaliumchlorid
kDa	Kilodalton

Kryo-EM	Kryo-Elektronenmikroskopie
L	Liter
LB	<i>Luria-Bertani</i>
M	Molar (mol/L)
m	milli
MALDI-TOF	Matrix Assisted Laser Deposition Ionization- Time Of Flight
MgCl ₂	Magnesiumchlorid
min	Minute
MMEJ	<i>Microhomology-mediated end joining</i>
MMR	<i>Mismatch-Reparatur</i>
MPD	2-Methyl-2,4-pentandiol
MS	Massenspektrum
MS/MS	Tandem Massenspektrometrie
n	nano
NaCl	Natriumchlorid
NaOH	Natriumhydroxid
NaSCN	Natriumthiocyanat
NER	Nukleotidexzisionsreparatur (<i>nucleotide excision repair</i>)
NHE	nicht-homologe Endverknüpfung
NHEJ	<i>Non homologous end joining</i>
Ni	Nickel
nm	Nanometer
NMR	<i>nuclear magnetic resonance</i> , Kernspinresonanz
NP-40	<i>NonidentTM P-40</i> , 4-Nonylphenyl-polyethylenglycol
nt	Nukleotid
NTA	Nitrilotriacetat
OB	Oligonukleotid/Oligosaccharid-bindendes Motiv
OBD	<i>origin binding motif</i>
OD ₆₀₀	Optische Dichte bei einer Wellenlänge von 600 nm
ODN	Oligodesoxynukleotid
ori	<i>origin of replication</i> , Replikationsursprung
ov	Überhang, <i>overhang</i>
PAA	Polyacrylamid
PAGE	Polyacrylamidgelelektrophorese

PAK	Polyzyklische aromatische Kohlenwasserstoffe
PCNA	<i>Proliferating-Cell-Nuclear-Antigen</i>
pH	Neg. dekadische Logarithmus der Wasserstoffionen-Konzentration
RFC	Replikationsfaktor C
RFU	<i>Relative Fluorescence Unit</i>
RNA	Ribonukleinsäure
RNase	Ribonuklease
RP	<i>Reversed phase</i>
RPA	Replikationsprotein A (<i>replication protein A</i>)
rpm	<i>rounds per minute</i> , Umdrehung pro Minute
RT	Raumtemperatur
s	Sekunden
SDS	Natriumdodecylsulfat
SEC	Größenausschluss-Chromatographie, <i>size exclusion column</i>
ss	Einzelsträngig
SSB	Einzelstrang-bindenden Proteinen
ssDNA	einzelsträngige DNA
T	Temperatur
<i>t</i>	<i>tert</i>
Tag	Helikase SV40 T Antigen
TBE	Tris/Borat/EDTA
TCR	Transkriptions-gekoppelte Reparatur
TEAA	Triethylammoniumacetat
TEMED	<i>N,N',N'',N'''</i> - Tetramethylethylendiamin
TFIIH	Transkriptionsfaktor II H
TLS	<i>Translesion DNA synthesis</i>
Tris	Tris-(hydroxymethyl)-aminomethan
TTD	Trichothiodystrophy
U	<i>Units</i>
UN	Ungeschädigt (<i>Undamaged</i>)
UNG	Uracil-DNA Glykosylase
UV	Ultraviolett
V	Volt
v/v	<i>volume per volume</i> , Volumenanteil

w/v	<i>weight per volume</i> , Massenanteil
wh	<i>winged helix</i>
XP	<i>Xeroderma Pigmentosum</i>
XPA-G	<i>Xeroderma Pigmentosum complementation group A-G</i>
ZnCl ₂	Zinkchlorid

8. Literaturverzeichnis

- [1] S. C. Koch, N. Simon, C. Ebert, T. Carell, Molecular mechanisms of xeroderma pigmentosum (XP) proteins, *Q. Rev. Biophys.*, **2016**, 49, e5.
- [2] E. Friedberg, G. Walker, W. Siede, P. v. d. Putte, DNA repair and mutagenesis, *Trends Biochem. Sci.*, **1995**, 20, 440.
- [3] S. R. Rajsiki, B. A. Jackson, J. K. Barton, DNA repair: models for damage and mismatch recognition, *Mutat. Res., Fundam. Mol. Mech. Mutagen.*, **2000**, 447, 49-72.
- [4] T. Lindahl, B. Nyberg, Rate of depurination of native deoxyribonucleic acid, *Biochemistry*, **1972**, 11, 3610-3618.
- [5] T. Lindahl, B. Nyberg, Heat-induced deamination of cytosine residues in deoxyribonucleic acid, *Biochemistry*, **1974**, 13, 3405-3410.
- [6] L. C. Gillet, O. D. Schärer, Molecular mechanisms of mammalian global genome nucleotide excision repair, *Chem. Rev.*, **2006**, 106, 253-276.
- [7] A. Sarasin, An overview of the mechanisms of mutagenesis and carcinogenesis, *Mutat. Res., Rev. Mutat. Res.*, **2003**, 544, 99-106.
- [8] W. Yang, R. Woodgate, What a difference a decade makes: insights into translesion DNA synthesis, *PNAS*, **2007**, 104, 15591-15598.
- [9] H. Ohmori, E. C. Friedberg, R. P. Fuchs, M. F. Goodman, F. Hanaoka, D. Hinkle, T. A. Kunkel, C. W. Lawrence, Z. Livneh, T. Nohmi, The Y-family of DNA polymerases, *Mol. Cell.*, **2001**, 8, 7-8.
- [10] O. D. Schärer, Chemistry and biology of DNA repair, *Angew. Chem., Int. Ed.*, **2003**, 42, 2946-2974.
- [11] S. Prakash, R. E. Johnson, L. Prakash, Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function, *Annu. Rev. Biochem.*, **2005**, 74, 317-353.
- [12] N. Sugitani, S. M. Shell, S. E. Soss, W. J. Chazin, Redefining the DNA-binding domain of human XPA, *J. Am. Chem. Soc.*, **2014**, 136, 10830-10833.
- [13] J. J. Truglio, D. L. Croteau, B. Van Houten, C. Kisker, Prokaryotic nucleotide excision repair: the UvrABC system, *Chem. Rev.*, **2006**, 106, 233-252.
- [14] S. M. Shell, Y. Zou, in *Molecular Mechanisms of Xeroderma Pigmentosum*, Springer, **2008**, pp. 103-112.
- [15] A. Gratchev, in *Molecular Mechanisms of Xeroderma Pigmentosum*, Springer, **2008**, pp. 113-119.
- [16] B. Singer, J. Kusmierek, Chemical mutagenesis, *Annu. Rev. Biochem.*, **1982**, 51, 655-691.
- [17] W. L. de Laat, N. G. Jaspers, J. H. Hoeijmakers, Molecular mechanism of nucleotide excision repair, *Genes Dev.*, **1999**, 13, 768-785.
- [18] K. Bekki, A. Toriba, N. Tang, T. Kameda, K. Hayakawa, Biological effects of polycyclic aromatic hydrocarbon derivatives, *J. UOEH*, **2013**, 35, 17-24.
- [19] F. A. Beland, F. F. Kadlubar, Formation and persistence of arylamine DNA adducts in vivo, *Environ. Health Perspect.*, **1985**, 62, 19.
- [20] J. S. Felton, M. G. Knize, N. H. Shen, B. D. Andresen, L. F. Bjeldanes, F. T. Hatch, Identification of the mutagens in cooked beef, *Environ. Health Perspect.*, **1986**, 67, 17.
- [21] M. Maclure, R. Katz, M. S. Bryant, P. L. Skipper, S. R. Tannenbaum, Elevated blood levels of carcinogens in passive smokers, *Am. J. Public Health*, **1989**, 79, 1381-1384.
- [22] A. H. Conney, Induction of microsomal enzymes by foreign chemicals and carcinogenesis by polycyclic aromatic hydrocarbons: GHA Clowes Memorial Lecture, *Cancer Res.*, **1982**, 42, 4875-4917.
- [23] R. G. Harvey, *Polycyclic aromatic hydrocarbons: chemistry and carcinogenicity*, CUP Archive, **1991**.
- [24] V. M. Arlt, M. Stiborová, C. J. Henderson, M. Thiemann, E. Frei, D. Aimová, R. Singh, G. G. da Costa, O. J. Schmitz, P. B. Farmer, Metabolic activation of benzo [a] pyrene in vitro by

- hepatic cytochrome P450 contrasts with detoxification in vivo: experiments with hepatic cytochrome P450 reductase null mice, *Carcinogenesis*, **2008**, *29*, 656-665.
- [25] N. E. Geacintov, M. Cosman, B. E. Hingerty, S. Amin, S. Broyde, D. J. Patel, NMR solution structures of stereoisomeric covalent polycyclic aromatic carcinogen– DNA adducts: principles, patterns, and diversity, *Chem. Res. Toxicol.*, **1997**, *10*, 111-146.
- [26] R. J. Turesky, Heterocyclic aromatic amine metabolism, DNA adduct formation, mutagenesis, and carcinogenesis, *Drug Metab. Rev.*, **2002**, *34*, 625-650.
- [27] R. H. Heflich, R. E. Neft, Genetic toxicity of 2-acetylaminofluorene, 2-aminofluorene and some of their metabolites and model metabolites, *Mutat. Res., Rev. Genet. Toxicol.*, **1994**, *318*, 73-174.
- [28] G. R. Hoffmann, R. P. Fuchs, Mechanisms of frameshift mutations: insight from aromatic amines, *Chem. Res. Toxicol.*, **1997**, *10*, 347-359.
- [29] D. M. Grant, N. C. Hughes, S. A. Janezic, G. H. Goodfellow, H. J. Chen, A. Gaedigk, L. Y. Violeta, R. Grewal, Human acetyltransferase polymorphisms, *Mutat. Res., Fundam. Mol. Mech. Mutagen.*, **1997**, *376*, 61-70.
- [30] M. Al-Zoughool, G. Talaska, 4-Aminobiphenyl N-glucuronidation by liver microsomes: optimization of the reaction conditions and characterization of the UDP-glucuronosyltransferase isoforms, *J. Appl. Toxicol.*, **2006**, *26*, 524-532.
- [31] F. F. Kadlubar, J. A. Miller, E. C. Miller, Hepatic microsomal N-glucuronidation and nucleic acid binding of N-hydroxy arylamines in relation to urinary bladder carcinogenesis, *Cancer Res.*, **1977**, *37*, 805-814.
- [32] H.-G. Neumann, The role of DNA damage in chemical carcinogenesis of aromatic amines, *J. Cancer Res. Clin. Oncol.*, **1986**, *112*, 100-106.
- [33] C. M. King, S. J. Land, R. F. Jones, M. Debiec-Rychter, M.-S. Lee, C. Y. Wang, Role of acetyltransferases in the metabolism and carcinogenicity of aromatic amines, *Mutat. Res., Fundam. Mol. Mech. Mutagen.*, **1997**, *376*, 123-128.
- [34] W. Xue, D. Warshawsky, Metabolic activation of polycyclic and heterocyclic aromatic hydrocarbons and DNA damage: a review, *Toxicol. Appl. Pharmacol.*, **2005**, *206*, 73-93.
- [35] D. Kim, F. P. Guengerich, Cytochrome P450 activation of arylamines and heterocyclic amines, *Annu. Rev. Pharmacol. Toxicol.*, **2005**, *45*, 27-49.
- [36] T. Shimada, N. Murayama, H. Yamazaki, K. Tanaka, S. Takenaka, M. Komori, D. Kim, F. P. Guengerich, Metabolic activation of polycyclic aromatic hydrocarbons and aryl and heterocyclic amines by human cytochromes P450 2A13 and 2A6, *Chem. Res. Toxicol.*, **2013**, *26*, 529-537.
- [37] G. L. Borosky, Ultimate carcinogenic metabolites from aromatic and heterocyclic aromatic amines: a computational study in relation to their mutagenic potency, *Chem. Res. Toxicol.*, **2007**, *20*, 171-180.
- [38] E. M. Bomhard, B. A. Herbold, Genotoxic activities of aniline and its metabolites and their relationship to the carcinogenicity of aniline in the spleen of rats, *Crit. Rev. Toxicol.*, **2005**, *35*, 783-835.
- [39] P. C. Howard, R. H. Heflich, F. E. Evans, F. A. Beland, Formation of DNA adducts in vitro and in *Salmonella typhimurium* upon metabolic reduction of the environmental mutagen 1-nitropyrene, *Cancer Res.*, **1983**, *43*, 2052-2058.
- [40] W. G. Humphreys, F. F. Kadlubar, F. P. Guengerich, Mechanism of C8 alkylation of guanine residues by activated arylamines: evidence for initial adduct formation at the N7 position, *PNAS*, **1992**, *89*, 8278-8282.
- [41] R. A. McClelland, A. Ahmad, A. P. Dicks, V. E. Licence, Spectroscopic characterization of the initial C8 intermediate in the reaction of the 2-fluorenylnitrenium ion with 2'-deoxyguanosine, *J. Am. Chem. Soc.*, **1999**, *121*, 3303-3310.

- [42] P. Belguise-Valladier, R. P. Fuchs, N-2-aminofluorene and N-2 acetylaminofluorene Adducts: The Local Sequence Context of an Adduct and its Chemical Structure Determine its Replication Properties, *J. Mol. Biol.*, **1995**, 249, 903-913.
- [43] L. M. Eckel, T. R. Krugh, 2-Aminofluorene modified DNA duplex exists in two interchangeable conformations, *Nat. Struct. Mol. Biol.*, **1994**, 1, 89-94.
- [44] L. C. Gillet, J. Alzeer, O. D. Schärer, Site-specific incorporation of N-(deoxyguanosin-8-yl)-2-acetylaminofluorene (dG-AAF) into oligonucleotides using modified 'ultra-mild' DNA synthesis, *Nucleic Acids Res.*, **2005**, 33, 1961-1969.
- [45] X. Veaute, R. P. Fuchs, Polymorphism in N-2-acetylaminofluorene induced DNA structure as revealed by DNase I footprinting, *Nucleic Acids Res.*, **1991**, 19, 5603-5606.
- [46] S. Dutta, Y. Li, D. Johnson, L. Dzantiev, C. C. Richardson, L. J. Romano, T. Ellenberger, Crystal structures of 2-acetylaminofluorene and 2-aminofluorene in complex with T7 DNA polymerase reveal mechanisms of mutagenesis, *Proc. Natl. Acad. Sci. U. S. A.*, **2004**, 101, 16186-16191.
- [47] F. A. Beland, K. L. Dooley, C. Jackson, Persistence of DNA adducts in rat liver and kidney after multiple doses of the carcinogen N-hydroxy-2-acetylaminofluorene, *Cancer Res.*, **1982**, 42, 1348-1354.
- [48] P. Howard, D. Casciano, F. Beland, J. Shaddock, The binding of N-hydroxy-2-acetylaminofluorene to DNA and repair of the adducts in primary rat hepatocyte cultures, *Carcinogenesis*, **1981**, 2, 97-102.
- [49] E. Kriek, Persistent binding of a new reaction product of the carcinogen N-hydroxy-N-2-acetylaminofluorene with guanine in rat liver DNA in vivo, *Cancer Res.*, **1972**, 32, 2042-2048.
- [50] P. P. Fu, Metabolism of nitro-polycyclic aromatic hydrocarbons, *Drug Metab. Rev.*, **1990**, 22, 209-268.
- [51] R. R. Bonala, M. C. Torres, S. Attaluri, C. R. Iden, F. Johnson, Incorporation of N-2-deoxyguanosine metabolic adducts of 2-aminonaphthalene and 2-aminofluorene into oligomeric DNA, *Chem. Res. Toxicol.*, **2005**, 18, 457-465.
- [52] T. Zaliznyak, R. Bonala, F. Johnson, C. de los Santos, Structure and Stability of Duplex DNA Containing the 3-(Deoxyguanosin-N2-yl)-2-acetylaminofluorene (dG(N2)-AAF) Lesion: A Bulky Adduct that Persists in Cellular DNA, *Chem. Res. Toxicol.*, **2006**, 19, 745-752.
- [53] N. Kitsera, K. Gasteiger, B. Luhnshof, J. Allgayer, B. Epe, T. Carell, A. Khobta, Cockayne syndrome: varied requirement of transcription-coupled nucleotide excision repair for the removal of three structurally different adducts from transcribed DNA, *PLoS One*, **2014**, 9, e94405.
- [54] M. S. Wold, Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism, *Annu. Rev. Biochem.*, **1997**, 66, 61-92.
- [55] G. G. Oakley, S. M. Patrick, Replication protein A: directing traffic at the intersection of replication and repair, *Front. Biosci.: a journal and virtual library*, **2010**, 15, 883.
- [56] R. Chen, M. S. Wold, Replication protein A: Single-stranded DNA's first responder, *BioEssays*, **2014**, 36, 1156-1161.
- [57] R. Chen, S. Subramanyam, A. H. Elcock, M. Spies, M. S. Wold, Dynamic binding of replication protein A is required for DNA repair, *Nucleic Acids Res.*, **2016**, 44, 5758-5772.
- [58] E. Bochkareva, S. Korolev, S. P. Lees-Miller, A. Bochkarev, Structure of the RPA trimerization core and its role in the multistep DNA-binding mechanism of RPA, *EMBO J.*, **2002**, 21, 1855-1863.
- [59] T. Liu, J. Huang, Replication protein A and more: single-stranded DNA-binding proteins in eukaryotic cells, *Acta Biochim. Biophys. Sin.*, **2016**.
- [60] S. K. Deng, H. Chen, L. S. Symington, Replication protein A prevents promiscuous annealing between short sequence homologies: implications for genome integrity, *BioEssays*, **2015**, 37, 305-313.

- [61] S. K. Binz, A. M. Sheehan, M. S. Wold, Replication protein A phosphorylation and the cellular response to DNA damage, *DNA repair*, **2004**, 3, 1015-1024.
- [62] E. Fanning, V. Klimovich, A. R. Nager, A dynamic model for replication protein A (RPA) function in DNA processing pathways, *Nucleic Acids Res.*, **2006**, 34, 4126-4137.
- [63] A. Bochkarev, E. Bochkareva, From RPA to BRCA2: lessons from single-stranded DNA binding by the OB-fold, *Curr. Opin. Struct. Biol.*, **2004**, 14, 36-42.
- [64] C. A. Brosey, M.-E. Chagot, M. Ehrhardt, D. I. Pretto, B. E. Weiner, W. J. Chazin, NMR analysis of the architecture and functional remodeling of a modular multidomain protein, RPA, *J. Am. Chem. Soc.*, **2009**, 131, 6346-6347.
- [65] S. J. Brill, S. Bastin-Shanower, Identification and characterization of the fourth single-stranded-DNA binding domain of replication protein A, *Mol. Cell. Biol.*, **1998**, 18, 7225-7234.
- [66] J. Chen, S. Le, A. Basu, W. J. Chazin, J. Yan, Mechanochemical regulations of RPA's binding to ssDNA, *Sci. Rep.*, **2015**, 5.
- [67] C. Kim, B. F. Paulus, M. S. Wold, Interactions of human replication protein A with oligonucleotides, *Biochemistry*, **1994**, 33, 14197-14206.
- [68] I. M. Wyka, K. Dhar, S. K. Binz, M. S. Wold, Replication protein A interactions with DNA: differential binding of the core domains and analysis of the DNA interaction surface, *Biochemistry*, **2003**, 42, 12909-12918.
- [69] A. I. Arunkumar, M. E. Stauffer, E. Bochkareva, A. Bochkarev, W. J. Chazin, Independent and coordinated functions of replication protein A tandem high affinity single-stranded DNA binding domains, *J. Biol. Chem.*, **2003**, 278, 41077-41082.
- [70] C. Kim, R. Snyder, M. Wold, Binding properties of replication protein A from human and yeast cells, *Mol. Cell. Biol.*, **1992**, 12, 3050-3059.
- [71] E. Carmichael, J. Roome, A. Wahl, Binding of a sequence-specific single-stranded DNA-binding factor to the simian virus 40 core origin inverted repeat domain is cell cycle regulated, *Mol. Cell. Biol.*, **1993**, 13, 408-420.
- [72] B. Nguyen, J. Sokoloski, R. Galletto, E. L. Elson, M. S. Wold, T. M. Lohman, Diffusion of human replication protein A along single-stranded DNA, *J. Mol. Biol.*, **2014**, 426, 3246-3261.
- [73] E. Bochkareva, V. Belegu, S. Korolev, A. Bochkarev, Structure of the major single-stranded DNA-binding domain of replication protein A suggests a dynamic mechanism for DNA binding, *EMBO J.*, **2001**, 20, 612-618.
- [74] L. Cai, M. Roginskaya, Y. Qu, Z. Yang, Y. Xu, Y. Zou, Structural characterization of human RPA sequential binding to single-stranded DNA using ssDNA as a molecular ruler, *Biochemistry*, **2007**, 46, 8226-8233.
- [75] S. A. Bastin-Shanower, S. J. Brill, Functional Analysis of the Four DNA Binding Domains of Replication Protein A THE ROLE OF RPA2 IN ssDNA BINDING, *J. Biol. Chem.*, **2001**, 276, 36446-36453.
- [76] A. Bochkarev, R. A. Pfuetzner, A. M. Edwards, L. Frappier, Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA, *Nature*, **1997**, 385, 176.
- [77] L. J. Blackwell, J. A. Borowiec, Human replication protein A binds single-stranded DNA in two distinct complexes, *Mol. Cell. Biol.*, **1994**, 14, 3993-4001.
- [78] C. A. Brosey, C. Yan, S. E. Tsutakawa, W. T. Heller, R. P. Rambo, J. A. Tainer, I. Ivanov, W. J. Chazin, A new structural framework for integrating replication protein A into DNA processing machinery, *Nucleic Acids Res.*, **2013**, 41, 2313-2327.
- [79] S. Kumaran, A. G. Kozlov, T. M. Lohman, *Saccharomyces cerevisiae* replication protein A binds to single-stranded DNA in multiple salt-dependent modes, *Biochemistry*, **2006**, 45, 11958-11973.
- [80] F. E. Kemmerich, P. Daldrop, C. Pinto, M. Levikova, P. Cejka, R. Seidel, Force regulated dynamics of RPA on a DNA fork, *Nucleic Acids Res.*, **2016**, 44, 5837-5848.
- [81] H. Chen, M. Lisby, L. S. Symington, RPA coordinates DNA end resection and prevents formation of DNA hairpins, *Mol. Cell.*, **2013**, 50, 589-600.

- [82] L. I. Toledo, M. Altmeyer, M.-B. Rask, C. Lukas, D. H. Larsen, L. K. Povlsen, S. Bekker-Jensen, N. Mailand, J. Bartek, J. Lukas, ATR prohibits replication catastrophe by preventing global exhaustion of RPA, *Cell*, **2013**, *155*, 1088-1103.
- [83] S. J. Haring, A. C. Mason, S. K. Binz, M. S. Wold, Cellular functions of human RPA1 Multiple roles of domains in replication, repair, and checkpoints, *J. Biol. Chem.*, **2008**, *283*, 19095-19111.
- [84] J. Smith, R. Rothstein, A mutation in the gene encoding the *Saccharomyces cerevisiae* single-stranded DNA-binding protein Rfa1 stimulates a RAD52-independent pathway for direct-repeat recombination, *Mol. Cell. Biol.*, **1995**, *15*, 1632-1641.
- [85] C. S. Hass, K. Lam, M. S. Wold, Repair-specific functions of replication protein A, *J. Biol. Chem.*, **2012**, *287*, 3908-3918.
- [86] H. Dou, C. Huang, M. Singh, P. B. Carpenter, E. T. Yeh, Regulation of DNA repair through deSUMOylation and SUMOylation of replication protein A complex, *Mol. Cell.*, **2010**, *39*, 333-345.
- [87] G. G. Oakley, S. M. Patrick, J. Yao, M. P. Carty, J. J. Turchi, K. Dixon, RPA Phosphorylation in Mitosis Alters DNA Binding and Protein– Protein Interactions, *Biochemistry*, **2003**, *42*, 3255-3264.
- [88] S. K. Binz, Y. Lao, D. F. Lowry, M. S. Wold, The Phosphorylation Domain of the 32-kDa Subunit of Replication Protein A (RPA) Modulates RPA-DNA Interactions EVIDENCE FOR AN INTERSUBUNIT INTERACTION, *J. Biol. Chem.*, **2003**, *278*, 35584-35591.
- [89] W. Shi, Z. Feng, J. Zhang, I. Gonzalez-Suarez, R. P. Vanderwaal, X. Wu, S. N. Powell, J. L. R. Roti, S. Gonzalo, J. Zhang, The role of RPA2 phosphorylation in homologous recombination in response to replication arrest, *Carcinogenesis*, **2010**, *31*, 994-1002.
- [90] T. Melendy, B. Stillman, An interaction between replication protein A and SV40 T antigen appears essential for primosome assembly during SV40 DNA replication, *J. Biol. Chem.*, **1993**, *268*, 3389-3395.
- [91] J. T. Yeeles, T. D. Deegan, A. Janska, A. Early, J. F. Diffley, Regulated eukaryotic DNA replication origin firing with purified proteins, *Nature*, **2015**, *519*, 431-435.
- [92] A. Yuzhakov, Z. Kelman, M. O'Donnell, Trading places on DNA—a three-point switch underlies primer handoff from primase to the replicative DNA polymerase, *Cell*, **1999**, *96*, 153-163.
- [93] P. A. Bullock, D. Simmons, The initiation of simian virus 40 DNA replication in vitro, *Crit. Rev. Biochem. Mol. Biol.*, **1997**, *32*, 503-568.
- [94] Z. He, B. T. Brinton, J. Greenblatt, J. A. Hassell, C. J. Ingles, The transactivator proteins VP16 and GAL4 bind replication factor A, *Cell*, **1993**, *73*, 1223-1232.
- [95] M. P. Longhese, P. Plevani, G. Lucchini, Replication factor A is required in vivo for DNA replication, repair, and recombination, *Mol. Cell. Biol.*, **1994**, *14*, 7884-7890.
- [96] S. Liu, Z. Xu, H. Leng, P. Zheng, J. Yang, K. Chen, J. Feng, Q. Li, RPA binds histone H3-H4 and functions in DNA replication–coupled nucleosome assembly, *Science*, **2017**, *355*, 415-420.
- [97] A. Aboussekhra, M. Biggerstaff, M. K. Shivji, J. A. Vilpo, V. Moncollin, V. N. Podust, M. Protić, U. Hübscher, J.-M. Egly, R. D. Wood, Mammalian DNA nucleotide excision repair reconstituted with purified protein components, *Cell*, **1995**, *80*, 859-868.
- [98] M. Saijo, A. Takedachi, K. Tanaka, Nucleotide excision repair by mutant xeroderma pigmentosum group A (XPA) proteins with deficiency in interaction with RPA, *J. Biol. Chem.*, **2011**, *286*, 5476-5483.
- [99] T. A. Nagelhus, T. Haug, K. K. Singh, K. F. Keshav, F. Skorpen, M. Otterlei, S. Bharati, T. Lindmo, S. Benichou, R. Benarous, A sequence in the N-terminal region of human uracil-DNA glycosylase with homology to XPA interacts with the C-terminal part of the 34-kDa subunit of replication protein A, *J. Biol. Chem.*, **1997**, *272*, 6561-6566.
- [100] Y.-L. Lin, M. K. Shivji, C. Chen, R. Kolodner, R. D. Wood, A. Dutta, The evolutionarily conserved zinc finger motif in the largest subunit of human replication protein A is required

- for DNA replication and mismatch repair but not for nucleotide excision repair, *J. Biol. Chem.*, **1998**, 273, 1453-1461.
- [101] D. D. Villarreal, K. Lee, A. Deem, E. Y. Shim, A. Malkova, S. E. Lee, Microhomology directs diverse DNA break repair pathways and chromosomal translocations, *PLoS Genet*, **2012**, 8, e1003026.
 - [102] L. Krejci, V. Altmannova, M. Spirek, X. Zhao, Homologous recombination and its regulation, *Nucleic Acids Res.*, **2012**, 40, 5795-5818.
 - [103] E. P. Mimitou, L. S. Symington, Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing, *Nature*, **2008**, 455, 770-774.
 - [104] B. O. Krogh, L. S. Symington, Recombination proteins in yeast, *Annu. Rev. Genet.*, **2004**, 38, 233-271.
 - [105] R. B. Jensen, A. Carreira, S. C. Kowalczykowski, Purified human BRCA2 stimulates RAD51-mediated recombination, *Nature*, **2010**, 467, 678-683.
 - [106] D. Jackson, K. Dhar, J. K. Wahl, M. S. Wold, G. E. Borgstahl, Analysis of the human replication protein A: Rad52 complex: evidence for crosstalk between RPA32, RPA70, Rad52 and DNA, *J. Mol. Biol.*, **2002**, 321, 133-148.
 - [107] N. Sugitani, W. J. Chazin, Characteristics and concepts of dynamic hub proteins in DNA processing machinery from studies of RPA, *Prog. Biophys. Mol. Biol.*, **2015**, 117, 206-211.
 - [108] X. Jiang, V. Klimovich, A. I. Arunkumar, E. B. Hysinger, Y. Wang, R. D. Ott, G. D. Guler, B. Weiner, W. J. Chazin, E. Fanning, Structural mechanism of RPA loading on DNA during activation of a simple pre-replication complex, *EMBO J.*, **2006**, 25, 5516-5526.
 - [109] K. Treuner, U. Ramsperger, R. Knippers, Replication protein A induces the unwinding of long double-stranded DNA regions, *J. Mol. Biol.*, **1996**, 259, 104-112.
 - [110] A. Prakash, G. E. Borgstahl, in *The Eukaryotic Replisome: a Guide to Protein Structure and Function*, Springer, **2012**, pp. 171-196.
 - [111] J. Witosch, E. Wolf, N. Mizuno, Architecture and ssDNA interaction of the Timeless-Tipin-RPA complex, *Nucleic Acids Res.*, **2014**, gku960.
 - [112] G. Mer, A. Bochkarev, R. Gupta, E. Bochkareva, L. Frappier, C. J. Ingles, A. M. Edwards, W. J. Chazin, Structural basis for the recognition of DNA repair proteins UNG2, XPA, and RAD52 by replication factor RPA, *Cell*, **2000**, 103, 449-456.
 - [113] G. Mer, A. Bochkarev, W. Chazin, A. Edwards, in *Cold Spring Harbor symposia on quantitative biology*, Vol. 65, Cold Spring Harbor Laboratory Press, **2000**, pp. 193-200.
 - [114] D. M. Jacobs, A. S. Lipton, N. G. Isern, G. W. Daughdrill, D. F. Lowry, X. Gomes, M. S. Wold, Human replication protein A: global fold of the N-terminal RPA-70 domain reveals a basic cleft and flexible C-terminal linker, *J. Biomol. NMR*, **1999**, 14, 321-331.
 - [115] A. Bochkarev, E. Bochkareva, L. Frappier, A. M. Edwards, The crystal structure of the complex of replication protein A subunits RPA32 and RPA14 reveals a mechanism for single-stranded DNA binding, *EMBO J.*, **1999**, 18, 4498-4504.
 - [116] X. Deng, J. E. Habel, V. Kabaleeswaran, E. H. Snell, M. S. Wold, G. E. Borgstahl, Structure of the full-length human RPA14/32 complex gives insights into the mechanism of DNA binding and complex formation, *J. Mol. Biol.*, **2007**, 374, 865-876.
 - [117] J. Fan, N. P. Pavletich, Structure and conformational change of a replication protein A heterotrimer bound to ssDNA, *Genes Dev.*, **2012**, 26, 2337-2347.
 - [118] N. Sugitani, R. M. Sivley, K. E. Perry, J. A. Capra, W. J. Chazin, XPA: A key scaffold for human nucleotide excision repair, *DNA repair*, **2016**, 44, 123-135.
 - [119] G. W. Daughdrill, G. W. Buchko, M. V. Botuyan, C. Arrowsmith, M. S. Wold, M. A. Kennedy, D. F. Lowry, Chemical shift changes provide evidence for overlapping single-stranded DNA- and XPA-binding sites on the 70 kDa subunit of human replication protein A, *Nucleic Acids Res.*, **2003**, 31, 4176-4183.

- [120] B. C. ScienceFeldes, D. Bonatto, Overview of xeroderma pigmentosum proteins architecture, mutations and post-translational modifications, *Mutat. Res., Rev. Mutat. Res.*, **2015**, 763, 306-320.
- [121] G. W. Buchko, G. W. Daughdrill, R. De Lorimier, B. Sudha Rao, N. G. Isern, J. M. Lingbeck, J.-S. Taylor, M. S. Wold, M. Gochin, L. D. Spicer, Interactions of human nucleotide excision repair protein XPA with DNA and RPA70ΔC327: Chemical shift mapping and 15N NMR relaxation studies, *Biochemistry*, **1999**, 38, 15116-15128.
- [122] T. Ikegami, I. Kuraoka, M. Saijo, N. Kodo, Y. Kyogoku, K. Morikawa, K. Tanaka, M. Shirakawa, Solution structure of the DNA-and RPA-binding domain of the human repair factor XPA, *Nat. Struct. Mol. Biol.*, **1998**, 5, 701-706.
- [123] L. Li, X. Lu, C. A. Peterson, R. J. Legerski, An interaction between the DNA repair factor XPA and replication protein A appears essential for nucleotide excision repair, *Mol. Cell. Biol.*, **1995**, 15, 5396-5402.
- [124] E. Fadda, Role of the XPA protein in the NER pathway: A perspective on the function of structural disorder in macromolecular assembly, *Comput. Struct. Biotechnol. J.*, **2016**, 14, 78-85.
- [125] S. M. Shell, W. J. Chazin, XPF-ERCC1: on the bubble, *Structure*, **2012**, 20, 566-568.
- [126] E. A. Maltseva, N. I. Rechkunova, I. O. Petrusheva, W. Vermeulen, O. D. Schärer, O. I. Lavrik, Crosslinking of nucleotide excision repair proteins with DNA containing photoreactive damages, *Bioorg. Chem.*, **2008**, 36, 77-84.
- [127] I. L. Hermanson-Miller, J. J. Turchi, Strand-specific binding of RPA and XPA to damaged duplex DNA, *Biochemistry*, **2002**, 41, 2402-2408.
- [128] Y. Liu, Y. Liu, Z. Yang, C. Utzat, G. Wang, A. K. Basu, Y. Zou, Cooperative interaction of human XPA stabilizes and enhances specific binding of XPA to DNA damage, *Biochemistry*, **2005**, 44, 7361-7368.
- [129] U. Camenisch, R. Dip, S. B. Schumacher, B. Schuler, H. Naegeli, Recognition of helical kinks by xeroderma pigmentosum group A protein triggers DNA excision repair, *Nat. Struct. Mol. Biol.*, **2006**, 13, 278-284.
- [130] M. Wakasugi, A. Sancar, Order of assembly of human DNA repair excision nuclease, *J. Biol. Chem.*, **1999**, 274, 18759-18768.
- [131] S. C. Shuck, E. A. Short, J. J. Turchi, Eukaryotic nucleotide excision repair: from understanding mechanisms to influencing biology, *Cell Res.*, **2008**, 18, 64-72.
- [132] S. M. Patrick, J. J. Turchi, Xeroderma pigmentosum complementation group A protein (XPA) modulates RPA-DNA interactions via enhanced complex stability and inhibition of strand separation activity, *J. Biol. Chem.*, **2002**, 277, 16096-16101.
- [133] M. Missura, T. Buterin, R. Hindges, U. Hübscher, J. Kaspáková, V. Brabec, H. Naegeli, Double-check probing of DNA bending and unwinding by XPA-RPA: an architectural function in DNA repair, *EMBO J.*, **2001**, 20, 3554-3564.
- [134] S. Ziani, Z. Nagy, S. Alekseev, E. Soutoglou, J.-M. Egly, F. Coin, Sequential and ordered assembly of a large DNA repair complex on undamaged chromatin, *J. Cell. Biol.*, **2014**, jcb.201403096.
- [135] C.-H. Park, D. Mu, J. T. Reardon, A. Sancar, The general transcription-repair factor TFIIH is recruited to the excision repair complex by the XPA protein independent of the TFIIIE transcription factor, *J. Biol. Chem.*, **1995**, 270, 4896-4902.
- [136] Y. S. Krasikova, N. Rechkunova, E. Maltseva, I. Petrusheva, V. Silnikov, T. Zatsepin, T. Oretskaya, O. Scharer, O. Lavrik, Interaction of nucleotide excision repair factors XPC-HR23B, XPA, and RPA with damaged DNA, *Biochemistry Moscow*, **2008**, 73, 886-896.
- [137] Z.-g. Yang, L. Y. Mao, Y. Zou, Y. Liu, J.-T. Zhang, Dimerization of human XPA and formation of XPA2-RPA protein complex, *Biochemistry*, **2002**, 41, 13012.
- [138] J.-S. You, M. Wang, S.-H. Lee, Biochemical analysis of the damage recognition process in nucleotide excision repair, *J. Biol. Chem.*, **2003**, 278, 7476-7485.

- [139] T. Mittag, L. E. Kay, J. D. Forman-Kay, Protein dynamics and conformational disorder in molecular recognition, *J. Mol. Recognit.*, **2010**, 23, 105-116.
- [140] J. Hu, Z. Hu, Y. Zhang, X. Gou, Y. Mu, L. Wang, X.-Q. Xie, Metal binding mediated conformational change of XPA protein: a potential cytotoxic mechanism of nickel in the nucleotide excision repair, *J. Mol. Model.*, **2016**, 22, 1-19.
- [141] A. O. Frank, B. Vangamudi, M. D. Feldkamp, E. M. Souza-Fagundes, J. W. Luzwick, D. Cortez, E. T. Olejniczak, A. G. Waterson, O. W. Rossanese, W. J. Chazin, Discovery of a potent stapled helix peptide that binds to the 70N domain of replication protein A, *J. Med. Chem.*, **2014**, 57, 2455-2461.
- [142] N. Simon, C. Ebert, S. Schneider, Structural Basis for Bulky-Adduct DNA-Lesion Recognition by the Nucleotide Excision Repair Protein Rad14, *Chem. - Eur. J.*, **2016**, 22, 10782-10785.
- [143] S. Shibutani, R. Gentles, F. Johnson, A. P. Grollman, Isolation and characterization of oligodeoxynucleotides containing dG-N2-AAF oxidation products of dG-C8-AAF, *Carcinogenesis*, **1991**, 12, 813-818.
- [144] E. Martin Armbrecht, Detektion von Kontaminationen in DNA und Protein-Proben durch photometrische Messungen, https://online-shop.eppendorf.de/DE-de/eshopdownload/downloadbykey/59828_1, **2013**, zuletzt aufgerufen am 20.04.2017.
- [145] B. Luhnshof, N. Kitsera, D. Warken, T. Lingg, B. Epe, A. Khobta, Generation of reporter plasmids containing defined base modifications in the DNA strand of choice, *Anal. Biochem.*, **2012**, 425, 47-53.
- [146] A. Khobta, S. Anderhub, N. Kitsera, B. Epe, Gene silencing induced by oxidative DNA base damage: association with local decrease of histone H4 acetylation in the promoter region, *Nucleic Acids Res.*, **2010**, 38, 4285-4295.
- [147] M. Lukin, C. de los Santos, NMR structures of damaged DNA, *Chem. Rev.*, **2006**, 106, 607-686.
- [148] Z. Yang, M. Roginskaya, L. C. Colis, A. K. Basu, S. M. Shell, Y. Liu, P. R. Musich, C. M. Harris, T. M. Harris, Y. Zou, Specific and Efficient Binding of Xeroderma Pigmentosum Complementation Group A to Double-Strand/Single-Strand DNA Junctions with 3'-and/or 5'-ssDNA Branches, *Biochemistry*, **2006**, 45, 15921-15930.
- [149] Thermo Scientific, PageRuler™ Unstained Protein Ladder, <https://www.thermofisher.com/order/catalog/product/26614>, **2015**, zuletzt aufgerufen am 19.04.2017.
- [150] NEB, 1 kb DNA Ladder, <https://www.neb.com/products/n3232-1-kb-dna-ladder>, zuletzt aufgerufen am 19.04.2017.
- [151] NEB, 100 bp DNA Ladder, <https://www.neb.com/products/n3231-100-bp-dna-ladder>, zuletzt aufgerufen am 19.04.2017.
- [152] iba, Expression and purification of proteins using 6xHistidine-tag, <https://www.iba-lifesciences.com/isotope/2/2-3201-010-Manual-6xHistidine-tag.pdf>, **2012**, zuletzt aufgerufen am 03.05.2017.
- [153] GE Healthcare Life Sciences, Ion Exchange Chromatography, Principles and Methods, http://www.google.de/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwjfsfuLndbTAhVHQBoKHeF5AR4QFggsMAA&url=http%3A%2F%2Fproteins.gelifesciences.com%2F%2Fmedia%2Fprotein-purification-ib%2Fdocuments%2Fhandbooks%2Fion_exchange_chromatography.pdf%3Ffla%3Den&usg=AFQjCNEY6sF611EQUyv7ERlk4tAxZhbNvg, **2016**, zuletzt aufgerufen am 03.05.2017.
- [154] GE Healthcare Life Sciences, <http://www.gelifesciences.com/webapp/wcs/stores/servlet/productById/en/GELifeSciences-de/17040601>, zuletzt aufgerufen am 03.05.2017.
- [155] GE Healthcare Life Sciences, Handbook: Size Exclusion Chromatography, Principles and Methods, <http://www.gelifesciences.com/webapp/wcs/stores/servlet/catalog/en/GELifeSciences-de/service-and-support/handbooks/>, **2014**, zuletzt aufgerufen am 03.05.2017.
- [156] L. A. Henricksen, C. B. Umbricht, M. S. Wold, Recombinant replication protein A: expression, complex formation, and functional characterization, *J. Biol. Chem.*, **1994**, 269, 11121-11132.

- [157] I. Lascu, H. Porumb, T. Porumb, I. Abrudan, C. Tarmure, I. Petrescu, E. Presecan, I. Proinov, M. Telia, Ion-exchange properties of Cibacron Blue 3G-A Sepharose (Blue Sepharose) and the interaction of proteins with Cibacron Blue 3G-A, *J. Chromatogr. A*, **1984**, 283, 199-210.
- [158] GE Healthcare Life Sciences, <http://www.gelifesciences.com/webapp/wcs/stores/servlet/productById/en/GELifeSciences-de/17041201>, zuletzt aufgerufen am 03.05.2017.
- [159] Bio-Rad, CHT™ Ceramic Hydroxyapatite, <http://www.bio-rad.com/webroot/web/pdf/lsr/literature/LIT611E.PDF>, zuletzt aufgerufen am 03.05.2017.
- [160] F. Lottspeich, H. Zorbas, Bioanalytik Spektrum Akademischer Verlag GmbH, *Heidelberg, Berlin*, **1998**.
- [161] ThermoFisher, <https://www.thermofisher.com/de/de/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/gel-shift-assays-ems.html>, zuletzt aufgerufen am 03.05.2017.
- [162] S. Boivin, S. Kozak, R. Meijers, Optimization of protein purification and characterization using ThermoFluor screens, *Protein Expression Purif.*, **2013**, 91, 192-206.
- [163] Greiner Bio-One, Advanced High-Throughput Platforms for Protein Crystallisation, https://www.gbo.com/fileadmin/user_upload/Downloads/Scientific_Publications/F073016_Forum_No._7_Protein_Crystallisation_E_Update_2015.pdf, **2015**, zuletzt aufgerufen am 03.05.2017.
- [164] Gatan, K2 Summit camera enables scientists to break the 3 Å barrier in structural biology using cryo-electron microscopy (cryo-EM), <http://www.gatan.com/company/news/k2-summit-camera-enables-scientists-break-3-%C3%A5-barrier-structural-biology-using-cryo>, **2015**, zuletzt aufgerufen am 03.05.2017.
- [165] CCMR, Negative stain EM, http://www.ccmr.cornell.edu/categoryuser-instrumentselectron-and-optical-microscopysingleparticle_cryoem/negative-stain-em/, zuletzt aufgerufen am 03.05.2017.
- [166] S. Beaucage, M. Caruthers, Deoxynucleoside phosphoramidites—a new class of key intermediates for deoxypolynucleotide synthesis, *Tetrahedron Lett.*, **1981**, 22, 1859-1862.
- [167] F. Johnson, C.-Y. Huang, P.-L. Yu, Synthetic and oxidative studies on 8-(arylamino)-2'-deoxyguanosine and-guanosine derivatives, *Environ. Health Perspect.*, **1994**, 102, 143.
- [168] O. Maillard, U. Camenisch, K. B. Blagoev, H. Naegeli, Versatile protection from mutagenic DNA lesions conferred by bipartite recognition in nucleotide excision repair, *Mutat. Res., Rev. Mutat. Res.*, **2008**, 658, 271-286.
- [169] L. M. Iakoucheva, A. L. Kimzey, C. D. Masselon, J. E. Bruce, E. C. Garner, C. J. Brown, A. K. Dunker, R. D. Smith, E. J. Ackerman, Identification of intrinsic order and disorder in the DNA repair protein XPA, *Protein Sci.*, **2001**, 10, 560-571.
- [170] L. M. Iakoucheva, A. L. Kimzey, C. D. Masselon, R. D. Smith, A. K. Dunker, E. J. Ackerman, Aberrant mobility phenomena of the DNA repair protein XPA, *Protein Sci.*, **2001**, 10, 1353-1362.
- [171] T. Hey, G. Lipps, G. Krauss, Binding of XPA and RPA to damaged DNA investigated by fluorescence anisotropy, *Biochemistry*, **2001**, 40, 2901-2910.
- [172] M. Götze, J. Pettelkau, R. Fritzsche, C. H. Ihling, M. Schäfer, A. Sinz, Automated assignment of MS/MS cleavable cross-links in protein 3D-structure analysis, *J. Am. Chem. Soc.*, **2015**, 26, 83-97.
- [173] M. Grimm, T. Zimniak, A. Kahraman, F. Herzog, xVis: a web server for the schematic visualization and interpretation of crosslink-derived spatial restraints, *Nucleic Acids Res.*, **2015**, gkv463.
- [174] J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized ppb-range mass accuracies and proteome-wide protein quantification, *Nat. Biotechnol.*, **2008**, 26, 1367-1372.